

UNIVERSIDADE FEDERAL DO PARANÁ

ANA PAULA PETRIU FERREIRA ENGELBERT

PRODUÇÃO E PERCEPÇÃO DA QUALIDADE DE VOZ EM PORTUGUÊS E
INGLÊS POR BRASILEIROS BILÍNGUES

CURITIBA

2015

ANA PAULA PETRIU FERREIRA ENGELBERT

PRODUÇÃO E PERCEPÇÃO DA QUALIDADE DE VOZ EM PORTUGUÊS E
INGLÊS POR BRASILEIROS BILÍNGUES

Tese apresentada ao curso de Pós-Graduação em Letras, Setor de Ciências Humanas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutor em Letras.

Orientadora: Prof^a. Dr^a. Adelaide Hercília Pescatori Silva

Co-orientadora: Prof^a. Dr^a. Jody Kreiman

CURITIBA

2015

Catálogo na publicação
Mariluci Zanela – CRB 9/1233
Biblioteca de Ciências Humanas e Educação - UFPR

Engelbert, Ana Paula Petriu Ferreira

Produção e percepção da qualidade de voz em português e inglês por
brasileiros bilíngues / Ana Paula Petriu Ferreira Engelbert – Curitiba, 2015.
182 f.

Orientadora: Profa. Dra. Adelaide Hercília Pescatori Silva

Co-orientadora: Profa. Dra. Jody Kreiman

Tese (Doutorado em Letras) – Setor de Ciências Humanas da
Universidade Federal do Paraná.

1. Absorção do som. 2. Voz – Qualidade. 3. Voz – Língua portuguesa.
4. Voz – Língua inglesa. 5. Fala. 5. Acústica – Análise. I. Título.

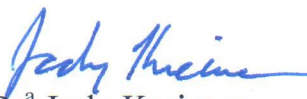
CDD 421.52




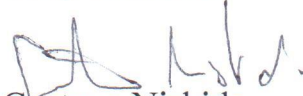
Setor de Ciências Humanas
Coordenação do Programa de Pós-Graduação em Letras
Tel./Fax: +55 41 3360-5102

Ata seiscentésima octogésima segunda, referente à sessão pública de defesa de tese para a obtenção de título de doutor a que se submeteu a doutoranda **ANA PAULA PETRIU FERREIRA ENGELBERT**. No dia nove de abril de dois mil e quinze, às quatorze horas, na sala 1013, 10.º andar, no Setor de Ciências Humanas da Universidade Federal do Paraná, foram instalados os trabalhos da Banca Examinadora, constituída pelos seguintes Professores Doutores: Adelaide Hercília Pescatori Silva, Presidente, Jody Kreiman (co-orientadora), Zuleica Camargo, Gustavo Nishida, Maria Lúcia de Castro Gomes e Denise Cristina Kluge designados pelo Colegiado do Curso de Pós-Graduação em Letras, para a sessão pública de defesa de dissertação intitulada **“PRODUÇÃO E PERCEPÇÃO DA QUALIDADE DE VOZ EM PORTUGUÊS E INGLÊS POR BRASILEIROS BILÍNGUES”**, apresentada por **ANA PAULA PETRIU FERREIRA ENGELBERT**. A sessão teve início com a apresentação oral da doutoranda sobre o estudo desenvolvido. Logo após, a senhora presidente dos trabalhos concedeu a palavra a cada um dos examinadores para as suas arguições. Em seguida, o candidato apresentou sua defesa. Na sequência, a Professora Adelaide Hercília Pescatori Silva retomou a palavra para as considerações finais. Na continuação, a Banca Examinadora, reunida sigilosamente, decidiu pela aprovação da candidata. Em seguida, a senhora Presidente declarou **APROVADA** a candidata, que recebeu o título de **Doutor em Letras**, área de concentração **Estudos Linguísticos**. A versão final da dissertação deverá ser encaminhada à Coordenação em até 60 dias. Encerrada a sessão, lavrou-se a presente ata, que vai assinada pela Banca Examinadora e pelo candidato. Feita em Curitiba, no dia nove de abril de dois mil e quinze.


Dr^a Adelaide Hercília Pescatori Silva


Dr^a Jody Kreiman


Dr^a Zuleica Camargo


Dr. Gustavo Nishida


Dr^a Maria Lúcia de Castro Gomes


Dr^a Denise Cristina Kluge


Ana Paula Petriu Ferreira Engelbert



Setor de Ciências Humanas
Coordenação do Programa de Pós-Graduação em Letras
Tel./Fax: +55 41 3360-5102

PARECER

Defesa de tese de doutorado de **ANA PAULA PETRIU FERREIRA ENGELBERT** para obtenção do título de **Doutor em Letras**.

Os abaixo-assinados Adelaide Hercília Pescatori Silva, Jody Kreiman, Zuleica Camargo, Gustavo Nishida, Maria Lúcia de Castro Gomes e Denise Cristina Kluge arguíram, nesta data, a candidata, a qual apresentou a tese “**PRODUÇÃO E PERCEPÇÃO DA QUALIDADE DE VOZ EM PORTUGUÊS E INGLÊS POR BRASILEIROS BILÍNGUES**”.

Procedida a arguição segundo o protocolo que foi aprovado pelo Colegiado do Curso, a Banca é de parecer que a candidata está apta ao título de **Doutor em Letras**, tendo merecido os conceitos abaixo:

Banca	Assinatura	APROVADO Não APROVADO
Dr ^a Adelaide Hercília Pescatori Silva (Presidente)		aprovada
Dr ^a Jody Kreiman (co-orientadora)		Aprovado
Dr ^a Zuleica Camargo		aprovada
Dr. Gustavo Nishida		Aprovado
Dr ^a Maria Lúcia de Castro Gomes		Aprovado
Dr ^a Denise Cristina Kluge		APROVADA

Curitiba, 09 de abril de 2015.

Prof^a Dr^a Maria José Foltran
Vice-Coordenadora

Maria José Foltran
Vice-Coordenadora
Matrícula SIAPE 0344084

Dedico esta tese a meu pai Acir (*in memoriam*), a minha mãe Clarisse, e ao Rodolfo, que são a razão da minha vida.

AGRADECIMENTOS

Primeiramente, agradeço a Deus por me guiar e me dar calma para seguir em frente com os meus objetivos e não desanimar com as dificuldades durante esses quatro anos.

Agradeço à minha família por sempre me incentivar e apoiar em todas as etapas da vida acadêmica. Ao meu pai, que partiu deste mundo antes de me ver alcançar o doutoramento, agradeço o amor e a confiança que sempre teve em mim. À minha mãe, que sempre esteve pronta a me ouvir nas horas de desespero e angústia, me acalmando e me incentivando a ter forças para vencer os obstáculos. Ao meu marido Rodolfo, que muito me apoiou ao longo de toda essa caminhada, me encorajando e me confortando com seu amor. Ao meu irmão Cesar, pelo carinho e prontidão a me ajudar sempre que necessário. A todos da família, que se acostumaram com a minha ausência nas reuniões familiares, mas que me apoiaram muito durante toda essa etapa. A vocês, o meu amor e minha gratidão.

Quero agradecer também à Universidade Federal do Paraná, instituição a que devo toda a minha formação. Presto meus agradecimentos a CAPES e à Fulbright pela bolsa de doutorado sanduíche que me possibilitou estudar nos Estados Unidos por nove meses. Agradeço à Universidade Tecnológica Federal do Paraná, em especial aos colegas do DALEM, pela oportunidade de me ausentar do trabalho para desenvolver e concluir a tese.

À minha orientadora, Adelaide, por ter aceitado me orientar no doutorado e acreditado no meu potencial. À minha orientadora nos Estados Unidos, Jody Kreiman, pela sua valiosa contribuição neste trabalho. A ela meu agradecimento pelo acolhimento na UCLA, por sua generosidade e pelo chá matinal de todos os dias. Quero agradecer também à Denise Kluge, pela preciosa ajuda em diversos aspectos deste trabalho, pela amizade, pelo carinho, pelas palavras de conforto. Sem vocês, não teria chegado até aqui.

Presto também meus agradecimentos aos colegas de trabalho da UTFPR, que sempre estiveram dispostos a me ajudar durante esses últimos quatro anos. Pelo carinho e incentivo da Ana Valéria, Miriam, Regina, Maristela, Silvana e Jacqueline. Pela troca de ideias, discussões e sugestões da Márcia, Malu e Andressa. A vocês, meu muito obrigada.

Agradeço aos participantes da banca pela disposição em avaliar esta tese e compartilhar do seu conhecimento. À Zuleica e à Denise, em especial, pelas valiosas sugestões na etapa da qualificação.

Por fim, aos participantes desta pesquisa, pela paciência e disponibilidade em fazer parte deste trabalho.

RESUMO

PRODUÇÃO E PERCEPÇÃO DA QUALIDADE DE VOZ EM PORTUGUÊS E INGLÊS POR BRASILEIROS BILÍNGUES

A qualidade de voz resulta da modulação dos articuladores do trato vocal e da configuração laríngea, produzindo um efeito de longo-termo na produção e percepção da fala. Bilíngues ou multilíngues podem alterar a qualidade de voz quando falam línguas diferentes, ou combinar qualidades das línguas que falam, conforme estudos na área. A motivação para tal alteração pode estar relacionada à personalidade, estado emocional e atitude do falante para com a L2 em questão. O presente estudo investiga a produção da voz e a percepção de sua qualidade em português brasileiro (PB) e inglês (IN) por falantes e ouvintes bilíngues brasileiros. Assim, os objetivos principais deste trabalho são: (a) descrever e comparar a produção de falantes brasileiros bilíngues em emissões no PB e no IN no que concerne às características vocais descritas por medidas acústicas; (b) examinar a percepção de qualidade de voz em PB e em IN de um mesmo falante por ouvintes leigos, também bilíngues; (c) averiguar possíveis relações entre as medidas acústicas de produção da voz e os julgamentos de percepção da qualidade de voz. Para atingir tais objetivos, foram realizados um experimento de produção e um de percepção. O experimento de produção consistiu na gravação de emissões em PB e em IN em tarefas de leitura de frases-veículo e de textos, bem como amostras em fala semiespontânea. Participaram desse experimento 16 brasileiros falantes nativos de PB e de IN como L2. As medidas acústicas utilizadas foram $H1^*-H2^*$, $H2^*-H4^*$, $H4^*-2\text{kHz}$, $2\text{kHz}-5\text{kHz}$, razão alfa, $D0225\text{kHz}$, $D2558\text{kHz}$, e média e extensão de f_0 . Os resultados mostram diferenças estatisticamente significativas entre médias e extensão de f_0 entre as línguas, bem como entre medidas de declínio espectral entre línguas e tarefas. O experimento de percepção consistiu na tarefa de discriminação entre emissões de fala em PB e em IN de um mesmo falante por outros 14 participantes bilíngues que não tinham experiência em avaliação de vozes. Foi investigado se esses ouvintes leigos percebiam possíveis diferenças de qualidade vocal entre as línguas, como eles quantificavam e como descreviam essas diferenças. A partir dos resultados dos dois experimentos, foram rodados testes estatísticos para verificar uma possível correlação entre diferenças de produção da voz e a percepção dessas diferenças pelos ouvintes leigos. De maneira geral, os ouvintes perceberam algumas diferenças entre as emissões em IN em relação ao PB e houve correlações entre os julgamentos perceptuais e as medidas acústicas realizadas nos dados de produção.

Palavras-chave: qualidade de voz; fala bilíngue; análise acústica.

ABSTRACT

VOICE QUALITY PRODUCTION AND PERCEPTION IN BRAZILIAN PORTUGUESE AND ENGLISH BY BRAZILIAN BILINGUALS

Voice quality results from the modulation of vocal tract articulators and laryngeal activity and produces a long-term effect in speech production and perception. Bilingual or multilingual speakers can alter voice quality when speaking different languages or even combine aspects from the languages they speak, according to evidences shown in research in the field. This change is usually related to the speaker's personality, emotional state and attitude towards the L2 and their speakers. Thus, the present study aims at investigating voice production and the perception of its quality in Brazilian Portuguese (BP) and English (EN) by Brazilian bilingual speakers and listeners. Thus, our aims were: (a) describe and compare Brazilian bilinguals' voice production in BP and EN through acoustic measures; (b) test Brazilian naïve listeners' perception of voice quality in same speaker/different language samples; (c) check for correlations between results from production and perception of voice quality. In order to achieve such goals, production and perception experiments were designed. The production experiment consisted of the recordings of reading and semi-spontaneous tasks in BP and EN. Sixteen Brazilian bilinguals participated in this experiment. The acoustic measures used to analyze production data were $H1^*-H2^*$, $H2^*-H4^*$, $H4^*-2\text{kHz}$, $2\text{kHz}-5\text{kHz}$, alpha ratio, $D0225\text{kHz}$, $D2558\text{kHz}$, and f_0 mean and range. Preliminary results show statistically significant differences between languages in f_0 measures, as well as in spectral characteristics between languages and tasks. The perception experiment consisted of a same speaker/different language discrimination task by Brazilian naïve listeners, also BP-EN bilinguals. The aim of this experiment was to find out whether listeners perceive possible differences in voice quality between BP and EN speech, how they quantify this difference and how they describe them. From the results of both experiments, it was possible to verify some correlations between differences in voice production and differences actually perceived by listeners.

Keywords: voice quality; L2 speech; acoustic analysis.

LISTA DE FIGURAS

FIGURA 1 - A PIRÂMIDE DA PROSÓDIA	19
FIGURA 2 - MÚSCULOS E LIGAMENTOS ENVOLVIDOS NA FONAÇÃO	34
FIGURA 3 - CONTÍNUO DE FONAÇÃO	36
FIGURA 4 - FORMA DA ONDA E ESPECTROGRAMA DA FONAÇÃO MODAL, SOPROSA E CREPITANTE, RESPECTIVAMENTE	41
FIGURA 5 - ESPECTROS FFT DA VOGAL /a/ NA VOZ MODAL, SOPROSA E CREPITANTE	42
FIGURA 6 - MODELO LF DA FONTE GLÓTICA	44
FIGURA 7 - MODELO FONTE-FILTRO DE PRODUÇÃO DA FALA	46
FIGURA 8 - ILUSTRAÇÃO ESQUEMÁTICA DO ESPECTRO ESPERADO PARA UM SINAL DE FALA	49
FIGURA 9 - REPRESENTAÇÃO DA TELA DA APRESENTAÇÃO DOS ESTÍMULOS NO PROGRAMA SKY	89
FIGURA 10 - REPRESENTAÇÃO DA TELA APÓS O JULGAMENTO DOS ESTÍMULOS PELOS OUVINTES	90
FIGURA 11 - MÉDIA E DP DOS JULGAMENTOS DE SOTAQUE ESTRANGEIRO	91
FIGURA 12 - CRITÉRIO DE SELEÇÃO DA PORÇÃO DA VOGAL PARA A ANÁLISE DA FONAÇÃO	94
FIGURA 13 - ETIQUETAGEM DA PORÇÃO DA VOGAL PARA A ANÁLISE DA FONAÇÃO	94
FIGURA 14 - TELA DO TP PARA O JULGAMENTO DAS VOZES EM “IGUAL” OU “DIFERENTE”	99
FIGURA 15 - TELA DO TP COM A ESCALA DESLIZANTE PARA A QUANTIFICAÇÃO DA DIFERENÇA ENTRE AS VOZES	100
FIGURA 16 - GRÁFICO DE CORRELAÇÃO ENTRE MÉDIAS DE F0 EM IN E IDADE DE AQUISIÇÃO	117

LISTA DE TABELAS

TABELA 1 -	JULGAMENTO DO SOTAQUE ESTRANGEIRO DOS PARTICIPANTES BRASILEIROS POR FALANTES NATIVOS DE IN	90
TABELA 2 -	IDADE, SEXO, ORIGEM, INÍCIO DE APRENDIZADO E TEMPO DE ESTUDO NO IN DOS OUVINTES	98
TABELA 3 -	EXPERIÊNCIA EM PAÍS DE LI, AUTOAVALIAÇÃO DE PROFICIÊNCIA E SOTAQUE	98
TABELA 4 -	MÉDIA E DP DAS MEDIDAS DE CURTO TERMO PARA OS PARTICIPANTES DO SEXO MASCULINO (N=6) E DO SEXO FEMININO (N=10), EM DECIBÉIS (dB)	105
TABELA 5 -	MÉDIA E DP DAS DO PPC PARA OS PARTICIPANTES DO SEXO MASCULINO (N=6) E DO SEXO FEMININO (N=10), EM Db	105
TABELA 6 -	MÉDIA E DP DAS MEDIDAS DE DECLÍNIO ESPECTRAL DE ELT DA LEITURA DE TEXTO DOS PARTICIPANTES DO SEXO MASCULINO (N=6) E DO FEMININO (N=10), EM dB.	106
TABELA 7 -	MÉDIA E DP DAS MEDIDAS DE DECLÍNIO ESPECTRAL DE ELT DA FALA SEMI-ESPONTÂNEA DOS PARTICIPANTES DO SEXO MASCULINO (N=6) E DO FEMININO (N=10), EM dB.	107
TABELA 8 -	MÉDIA E DP DAS MEDIDAS DE F0 DA LEITURA DE TEXTOS EM PB E IN POR PARTICIPANTES HOMENS (N=6) E MULHERES (N=10), EM Hz	108
TABELA 9 -	MÉDIA E DP DAS MEDIDAS DE F0 DA FALA SEMI-ESPONTÂNEA EM PB E IN POR PARTICIPANTES HOMENS (N=6) E MULHERES (N=10), EM Hz	109
TABELA 10 -	RESULTADOS DO TESTE WILCOXON ENTRE PB E IN PARA MEDIDAS DE CURTO TERMO	109
TABELA 11 -	RESULTADOS DO TESTE WILCOXON ENTRE PB E IN PARA MEDIDAS ESPECTRAIS DE CURTO TERMO	110
TABELA 12 -	RESULTADOS DO TESTE WILCOXON PARA MEDIDAS DE ELT ENTRE TAREFAS LEITURA DE TEXTO (TX) E FALA SEMIESPONTÂNEA (SE)	112
TABELA 13 -	RESULTADOS DO TESTE WILCOXON ENTRE PB E IN PARA MEDIDAS DE F0	113
TABELA 14 -	RESULTADOS DO TESTE WILCOXON PARA MEDIDAS DE F0 DO ENTRE AS TAREFAS LEITURA DE TEXTO (TX) E DE FALA SEMI-ESPONTÂNEA (SE)	113

TABELA 15 -	RESULTADOS DO TESTE MANN WHITNEY ENTRE HOMENS E MULHERES PARA AS MEDIDAS ESPECTRAIS DE CURTO E LONGO TERMO	114
TABELA 16 -	RESULTADOS DO TESTE MANN WHITNEY ENTRE HOMENS E MULHERES PARA AS MEDIDAS DE F0 NA LEITURA DE TEXTO (TX)	115
TABELA 17 -	RESULTADOS DO TESTE MANN WHITNEY ENTRE HOMENS E MULHERES PARA AS MEDIDAS DE F0 NA FALA SEMIESPONTÂNEA (SE)	115
TABELA 18 -	RESULTADOS DO TESTE PEARSON ENTRE A IDADE DE AQUISIÇÃO E AS MEDIDAS ACÚSTICAS	116
TABELA 19 -	MÉDIA E DP DOS JULGAMENTOS DE DIFERENÇAS ENTRE VOZES EM PB E EM IN PELOS OUVINTES	118
TABELA 20 -	MÉDIA E DP DOS JULGAMENTOS DE DIFERENÇAS ENTRE VOZES EM PB E EM IN PELOS OUVINTES QUE MAIS OBTIVERAM CORRELAÇÕES	119
TABELA 21 -	RESULTADOS DA ANÁLISE AGLOMERATIVA DE K-MÉDIAS PARA O AGRUPAMENTO DOS FALANTES	120
TABELA 22 -	RESULTADOS DO TESTE ANOVA SOBRE A ANÁLISE AGLOMERATIVA DE K-MÉDIAS	120
TABELA 23 -	RESULTADOS DO TESTE MANN WHITNEY PARA DIFERENÇAS ENTRE OS DOIS GRUPOS DE FALANTES	121
TABELA 24 -	MÉDIA E DP DAS MEDIDAS DE LONGO TERMO E DE F0 DA FALA SEMIESPONTÂNEA	122
TABELA 25 -	MÉDIA E DP DAS MEDIDAS DE ELT E F0 NA FALA SEMIESPONTÂNEA PARA OS HOMENS DOS GRUPOS 1 E 2 ..	123
TABELA 26 -	MÉDIA E DP DAS MEDIDAS DE ELT E F0 NA FALA SEMIESPONTÂNEA PARA AS MULHERES DOS GRUPOS 1 E 2	123

LISTA DE QUADROS

QUADRO 1 -	TIPOS DE JULGAMENTO DA VOZ	27
QUADRO 2 -	PESQUISAS PARA A CONSTRUÇÃO DO MODELO PSICOACÚSTICO DE PRODUÇÃO E PERCEPÇÃO DA VOZ	56
QUADRO 3 -	EVIDÊNCIAS PARA A CONSTRUÇÃO DO MODELO PSICOACÚSTICO DE PRODUÇÃO E PERCEPÇÃO DA VOZ	57
QUADRO 4 -	SEXO, IDADE E ORIGEM DOS PARTICIPANTES DO EXPERIMENTO DE PRODUÇÃO	86
QUADRO 5 -	DETALHES SOBRE A AQUISIÇÃO E EXPERIÊNCIA NO IN DOS PARTICIPANTES DO EXPERIMENTO DE PRODUÇÃO	86
QUADRO 6 -	DESCRIÇÕES DAS VOZES EM PB E IN DOS FALANTES DO GRUPO 1	125
QUADRO 7 -	RESUMO DAS PERGUNTAS DE PESQUISA, HIPÓTESES E RESULTADOS	136

LISTA DE ABREVIATURAS

L1-	Língua Materna
L2 -	Língua Estrangeira / Segunda Língua
PB-	Português Brasileiro
IN-	Inglês
SLM-	<i>Speech Learning Model</i>
PAM-	<i>Perceptual Assimilation Model</i>
ELT -	Espectro de Longo Termo
MDS-	<i>Multidimensional Scaling</i>
FFT-	<i>Fast Fourier Transform</i>
LF-	<i>Liljencrants-Fant Model</i>
dB-	Decibel
Hz-	Hertz
DP-	Desvio Padrão
SE-	Semiespontânea (fala)
TX-	Texto
PPC-	Pico de Proeminência Cepstral

SUMÁRIO

1 INTRODUÇÃO	1
1.2 OBJETIVOS	6
1.3 PERGUNTAS DE PESQUISA E HIPÓTESES	7
1.4 ORGANIZAÇÃO DA TESE	10
2 PROSÓDIA	12
2.1 CARACTERÍSTICAS PROSÓDICAS DAS LÍNGUAS	15
2.1.1 <i>Pitch</i> e frequência fundamental	15
2.1.2 Volume, intensidade e amplitude	16
2.1.3 Duração e pistas temporais	17
2.1.4 Qualidade de voz	18
2.2 NÍVEIS FUNCIONAIS DAS PISTAS PROSÓDICAS	18
2.3 QUALIDADE DE VOZ E A PRAGMÁTICA	20
2.4 PROSÓDIA E ASPECTOS PARLINGUÍSTICOS	21
2.5 CONCLUSÃO DO CAPÍTULO	26
3 QUALIDADE DE VOZ	27
3.1 DEFINIÇÃO DE QUALIDADE DE VOZ	28
3.2 COMO MEDIR A VOZ E SUA QUALIDADE	30
3.3 ANATOMIA E FUNCIONAMENTO DAS PREGAS VOCAIS	33
3.4 FONAÇÃO	35
3.4.1 Tipos de fonação	36
3.4.1.1 Fonação modal	37
3.4.1.2 Fonação não-modal	37
3.4.1.2.1 Voz crepitante	38
3.4.1.2.2 Voz soprosa	39
3.4.1.3 Fonação não-modal alofônica	39
3.4.2 Características acústicas dos tipos de fonação	40
3.4.3 Análise dos tipos de fonação	44

3.4.4 Medidas acústicas do domínio espectral	47
3.5 MEDIDAS DE F0	53
3.6 UM MODELO PSICOACÚSTICO PARA A ANÁLISE DA VOZ E SUA QUALIDADE.....	54
3.7 CONCLUSÃO DO CAPÍTULO	59
4 AQUISIÇÃO DAS CARACTERÍSTICAS PROSÓDICAS DA L2	60
4.1 FATORES QUE INFLUENCIAM A AQUISIÇÃO DE PADRÕES PROSÓDICOS DA L2	63
4.1.1 Idade de aquisição	63
4.1.2 Experiência linguística.....	64
4.1.3 Ambiente de aprendizagem.....	65
4.1.4 Fatores sociais e afetivos	66
4.2 QUALIDADE DE VOZ E AQUISIÇÃO DE L2	67
4.2.1 Estudos sobre a produção da voz por bilíngues.....	69
4.3 PERCEPÇÃO DA QUALIDADE DE VOZ EM L2.....	73
4.4 CONCLUSÃO DO CAPÍTULO	77
5 METODOLOGIA	79
5.1 ESTUDO PILOTO	79
5.2 EXPERIMENTO DE PRODUÇÃO.....	82
5.2.1 <i>Corpus</i>	82
5.2.1.1 Corpus para medidas espectrais de curto termo	82
5.2.1.2 Corpus para medidas de f0 e de ELT	84
5.2.2 Entrevista	84
5.2.3 Participantes.....	85
5.2.3.1 Julgamento do grau de sotaque dos falantes	87
5.2.4 Gravações	91
5.2.5 Realização das medidas acústicas.....	93
5.2.5.1 Medidas espectrais de curto termo.....	93
5.2.5.2 Medidas de f0 e ELT	95
5.3 EXPERIMENTO DE PERCEPÇÃO	97

5.3.1 Os ouvintes	97
5.3.2 Os estímulos	98
5.3.3 As instruções e a tarefa	99
5.4 DESCRIÇÃO E ANÁLISE DOS DADOS	101
5.5 CONCLUSÃO DO CAPÍTULO	102
6 RESULTADOS	103
6.1 EXPERIMENTO DE PRODUÇÃO	103
6.1.1 Agrupamento dos dados e estatística descritiva	103
6.1.1.2 Medidas espectrais de curto termo	104
6.1.1.3 Medidas espectrais de longo termo	106
6.1.1.4 Medidas de f_0	107
6.1.2 Testagem das hipóteses	109
6.1.2.1 Hipótese 1	109
6.1.2.2 Hipóteses 2.1 e 2.2	110
6.1.2.3 Hipóteses 3.1 e 3.2	113
6.1.2.4 Hipótese 4	114
6.1.2.5 Hipótese 5	116
6.2 EXPERIMENTO DE PERCEPÇÃO	117
6.3 RELAÇÃO ENTRE PRODUÇÃO E PERCEPÇÃO	121
6.4 DISCUSSÃO	126
7. CONCLUSÃO	135
REFERÊNCIAS	141
ANEXOS	156

1 INTRODUÇÃO

Já ouvi muitas vezes que soo diferente quando falo inglês (IN)¹ - minha primeira língua estrangeira (L2) - do que quando falo português brasileiro (PB) - minha língua materna (L1). Também já ouvi outras pessoas que falam IN ou outra L2 relatarem a mesma experiência. Por conta dessa impressão, sempre presto atenção nas vozes das pessoas quando usam uma L2 em comparação com quando usam suas línguas maternas. Essa impressão me motivou a pesquisar a qualidade de voz na fala de bilíngues, que é o tema central desta tese. Segundo Esling (2013), pelo fato de a qualidade de voz ser socialmente indicial, indivíduos multilíngues mudam os ajustes na voz quando falam línguas diferentes ou mesmo combinam ajustes das línguas que falam. Assim, o objetivo principal desta tese é verificar se há mudanças na produção da voz e na percepção da qualidade da voz de bilíngues brasileiros quando falam português em comparação com quando falam IN como L2; se essa impressão de mudança realmente se confirma.

Não faz parte do escopo do presente trabalho discutir em profundidade o conceito ou a definição de bilinguismo, mas faz-se necessário justificar o uso do termo 'bilíngue' para designar o perfil do participante desta pesquisa. Há na literatura uma série de definições de bilinguismo que, segundo de Bot *et al.* (2005, p. 5), estão entre duas categorias extremas: a que considera bilíngue somente a pessoa que cresceu falando duas línguas (bilíngue real²) e tem o comando completo nas duas; e a segunda, que considera bilíngue aquele que tem qualquer conhecimento de outra língua que não a materna. Esses autores acreditam que a definição ideal seja algo entre esses dois extremos. Ademais, há uma vasta discussão em torno do que seja 'ter comando completo' em uma língua, ou mesmo se há pessoas verdadeiramente monolíngues, tornando a definição de bilinguismo ainda mais problemática. A definição vigente de bilinguismo (e

¹ Neste trabalho, optou-se pela variedade estadunidense do inglês por dois motivos: os dados do experimento de produção foram coletados nos Estados Unidos, e pelo fato de essa variedade ser a mais ensinada em escolas de idiomas no Brasil. Portanto, quando se utilizar a sigla IN, leia-se inglês da variedade estadunidense.

² Do inglês "*real bilingual*", aquele que adquire/aprende duas línguas simultaneamente.

também de multilinguismo) mais adotada, como afirmam Zimmer e Alves (2014), é a de que haja um *continuum* entre o monolinguismo e o bi/multilinguismo que contemple vários estágios de proficiência e ativação das línguas do falante. Essa proposta, que é de Grosjean (2008), tem como referencial o uso da(s) língua(s) pelo falante em diferentes situações comunicativas, de modo que o bilíngue utilize línguas diferentes para alcançar objetivos comunicativos em contextos sociolinguísticos diversos. A principal motivação pelo uso do termo ‘bilíngue’ na presente pesquisa, ao invés de utilizar o termo ‘aprendiz de L2’, vem do fato de que a maioria dos trabalhos sobre a produção e percepção da qualidade de voz adota esse termo para caracterizar os participantes falantes de duas línguas. Tais estudos são descritos na seção 4.2.

A definição dos termos voz e qualidade de voz também não é consensual, como será discutido no capítulo 3, mas, em linhas gerais, referem-se a características resultantes da modulação dos articuladores do trato vocal que produzem um efeito de longo-termo na produção da fala e sua percepção pelo ouvinte. Por meio da configuração do trato vocal e, conseqüentemente, da produção de diferentes características vocais, é possível ajustar a voz a determinadas situações. Isso acontece, pois há muitas outras informações na fala para além do conteúdo linguístico que se emprega durante a comunicação. Como falantes, fornecemos ao nosso ouvinte informações sobre quem somos, de onde viemos e o que estamos sentindo. Como ouvintes, também somos capazes de inferir uma série dessas características a partir da voz do falante. A qualidade vocal também é determinada pela composição fisiológica do trato, já que o formato e posição dos articuladores como os lábios, língua, mandíbula influenciam diretamente a produção dos sons da fala.

A qualidade de voz pode ter tanto função linguística, ou seja, que implique contrastes em uma dada língua, como pode ter função paralinguística, quando fornece informações físicas e psicossociais do falante. Neste trabalho, serão focalizadas as características paralinguísticas da qualidade vocal, já que não há no PB e nem no IN, a utilização de diferentes qualidades de voz que forneçam contraste linguístico.

Vale ressaltar que o tema deste trabalho não tem relação direta com o “sotaque” que uma pessoa pode ter quando fala uma L2 – característica comum a muitos bilíngues – mas sim com a produção da voz e com a percepção de sua qualidade. O sotaque é definido por Munro (1998, p. 139) como “a fala não-patológica produzida por aprendizes de segunda língua que difere de maneira parcialmente sistemática da fala característica dos nativos de um dado dialeto”³. Ou seja, o sotaque estrangeiro é comumente percebido pelos “desvios” fonético-fonológicos da produção de um falante não-nativo da L2 em relação aos padrões nativos dessa língua. Esses desvios podem acontecer tanto no domínio segmental quanto no prosódico, afetando a produção e a percepção dos contrastes linguísticos da L2. O presente estudo também não se propõe a tratar pontualmente de questões de inteligibilidade em L2. A inteligibilidade, segundo Munro e Derwing (1995), tem relação com a percepção do quanto é entendido de fato de um enunciado⁴. Nesse sentido, não temos o objetivo de verificar a inteligibilidade da produção dos participantes brasileiros bilíngues, embora acreditemos que os aspectos prosódicos que operam nas línguas têm um papel crucial na inteligibilidade da fala. Portanto, o presente trabalho verifica possíveis mudanças na produção da voz e na percepção da qualidade de voz por bilíngues no domínio paralinguístico, e não terá como objetivo tratar de questões relacionadas à percepção e produção de sotaque estrangeiro e à inteligibilidade na L2.

O processo de aquisição⁵ de aspectos fonético-fonológicos do IN como L2 por aprendizes brasileiros tem sido amplamente investigado, sobretudo nas três últimas décadas. No que concerne à aquisição de características segmentais da L2, um número grande de pesquisas vem sendo realizado para caracterizar a fala do aprendiz estrangeiro e relacionar fatores linguísticos e extralinguísticos que influenciam a aprendizagem dos indivíduos. A concepção do Modelo de Aprendizagem da Fala (*Speech Learning Model* – SLM) por Flege (1995) contribuiu

³ As traduções dos textos cujos originais estão em IN são responsabilidade da autora deste trabalho. Os originais aparecerão em itálico, em notas de rodapé.

...non-pathological speech produced by second language learners that differs in partially systematic ways from the speech characteristic of native speakers of a given dialect

⁴ Para mais informações sobre inteligibilidade e aquisição de L2, ver Becker (2013).

⁵ Neste trabalho, não serão feitas distinções entre os termos “aquisição” e “aprendizagem” de L2.

sobremaneira para o desenvolvimento da pesquisa em descrição e análise da produção da fala em L2. Tal modelo tem como premissa que quanto maior a diferença fonética entre um som da L2 do som mais próximo da L1 do aprendiz, maior a probabilidade do indivíduo discernir os dois sons e, conseqüentemente, produzir e percebê-los com mais sucesso. Além da noção de similaridade fonética, o modelo prevê a criação de novas categorias a partir da percepção e, assim, a aquisição de contrastes fonêmicos da L2, principalmente por aprendizes em níveis mais proficientes. Já para análise da percepção da fala não-nativa, o modelo mais utilizado é o de Best e Tyler (2007), intitulado Modelo de Assimilação Perceptual (*Perceptual Assimilation Model* – PAM⁶). Esse modelo tem como principal premissa que a variação entre a discriminação de contrastes não-nativos se dá pela aptidão fonética dos sons a se encaixar em categorias da língua-alvo. Os dois modelos explicam a aquisição de fatores fonético-fonológicos de uma L2 pela interação existente entre os níveis segmentais da L1 e L2. Embora a presente pesquisa trate da fala bilíngue, não se pretende descrever ou analisar a produção e percepção dos brasileiros no que concerne aos contrastes do IN como L2, já que o domínio que se está investigando é o paralinguístico.

Um grande número de pesquisas sobre a aquisição do IN por brasileiros tem apresentado respostas a questões relacionadas ao processo de ensino/aprendizagem de características fonético-fonológicas em nível segmental, tanto no que concerne à produção quanto à percepção (Baptista e Silva Filho, 1997; Baptista, 2000; Koerich, 2002, 2006; Kluge, 2004, 2009; Rauber, 2006; Nobre-Oliveira, 2007; Silveira, 2004; Becker, 2007; Ferreira, 2007; Gomes, 2009). A aquisição de aspectos suprasegmentais de IN como L2 por brasileiros tem sido pouco estudada, salvo em trabalhos de Brawerman (2006) e Brawerman-Albini (2012), que investigaram a percepção e produção do acento lexical, bem como o efeito do seu treinamento; Mauad (2007), que verificou a produção de padrões entoacionais realizados por falantes nativos do IN e por aprendizes brasileiros dessa língua; Marta (2011), que comparou a produção e a percepção de curvas melódicas de enunciados declarativos e interrogativo-negativos do IN; e Rocca

⁶ O PAM foi apresentado em Best (1995), mas essa primeira versão do modelo era voltada para a aquisição de primeira língua. Posteriormente, em Best e Tyler (2007), o modelo foi proposto para a percepção na aquisição da L2.

(2011), que estudou o efeito do treinamento de aspectos da entoação do inglês pela utilização de instrumentos de análise acústica.

A carência de atenção dada aos aspectos suprasegmentais também se revela nos materiais de ensino de pronúncia do IN, sendo que a maioria traz atividades dedutivas e baseadas em regras. Jones e Evans (1995, p. 244) afirmam que é muito difícil encontrar materiais que "tentem vincular padrões rítmicos ao discurso, grau de formalidade, ou o estado emocional do falante."⁷ Porque pouca atenção tem sido dada ao uso da qualidade de voz em L2, pouco se sabe sobre como as pessoas mudam e o que difere quando eles falam uma língua estrangeira.

Quanto à pesquisa sobre a qualidade de voz no Brasil, o foco tem sido nas alterações na qualidade em casos clínicos (Camargo, 1996, 2002; Camargo *et al.* 2003; Camargo *et al.*, 2004; Lima 2008, 2012; Oliveira, 2011; Pessoa *et al.*, 2012), e no uso da voz e expressividade (Madureira, 2004; Viola e Madureira, 2007; Madureira e Camargo, 2010; Fontes e Madureira, 2014). Além dessas áreas, há estudos que têm se disposto a discutir a relação entre métodos perceptuais de avaliação da qualidade da voz e correlatos acústicos da fala (Rusilo *et al.*, 2011; Camargo *et al.*, 2012). Com tema mais próximo do tema do presente trabalho, o estudo de Camargo *et al.* (2013) visou verificar as diferentes qualidades de voz produzidas por um falante multilíngue em IN, espanhol e português. Nessa pesquisa, utilizou-se a avaliação perceptiva e a instrumental, esta por medidas acústicas, para descrever as alterações de qualidade de voz nas diferentes línguas do falante.

Utilizar-se da correlação entre as duas abordagens – a perceptiva e a instrumental – para tentar descrever e explicar a relação entre a percepção e a produção da qualidade voz tem sido a alternativa mais recorrente em estudos da área. Kent e Ball (2000) afirmam que, mesmo que métodos instrumentais sejam quantitativos e mais objetivos, não se sustentam sozinhos – o julgamento perceptivo-auditivo é necessário para a identificação de uma qualidade de voz, de maneira que as abordagens perceptiva e instrumental são complementares.

Assim, a intenção deste trabalho é comparar a qualidade vocal na produção de brasileiros quando falam sua língua materna, o PB, e o IN como L2, e na

⁷ ...link rhythmic patterns to discourse, degree of formality, or the emotional state of the speaker.

percepção da qualidade de voz por ouvintes leigos, a fim de verificar se realmente ocorrem mudanças ao falarmos outra língua e se tal mudança é percebida pelos ouvintes. Desta forma, pela escassez de trabalhos relacionados a questões prosódicas interlinguísticas, justifica-se a realização desta pesquisa, que pretende comparar a produção da fala em L1 e L2 em relação à qualidade de voz em dados intrafalantes pela análise perceptiva e acústica, a fim de entender melhor o que muda quando falamos línguas diferentes.

1.2 OBJETIVOS

O objetivo principal desta pesquisa é investigar a produção e a percepção da qualidade de voz em falantes bilíngues. De acordo com alguns estudos sobre o tema (Bruyninckx *et al.*, 1991; 1994; Harmegnies *et al.*, 1991; Todaka, 1995; Altenberg e Ferrand, 2006; Ng *et al.*, 2012; Camargo *et al.*, 2013), falantes bilíngues modificam aspectos da qualidade de voz quando falam línguas diferentes. Segundo essas pesquisas, os bilíngues realizam as mudanças pelo uso de diferentes tipos de fonação e diferenças médias e extensões de frequência fundamental. Além disso, o tipo de tarefa do experimento - leitura e fala semiespontânea ou espontânea - pode influenciar na produção da qualidade de voz.

Este estudo pretende descrever a produção da voz por brasileiros, falantes bilíngues de português como L1 e IN como L2, e comparar as amostras de fala de um mesmo falante nessas duas línguas. Para que se alcançar tal objetivo, foi realizado um teste de produção com a participação de 16 brasileiros, falantes bilíngues, na realização de tarefas de leitura e fala semiespontânea. Para realizar tal comparação, foram utilizadas medidas acústicas que refletem modificações na fonte glótica.

Outro objetivo desta pesquisa é o de investigar se eventuais mudanças na qualidade de voz produzidas por falantes bilíngues são percebidas por ouvintes leigos, também bilíngues. Para se atingir esse objetivo, foi realizado um teste de percepção, por meio de uma tarefa de discriminação, em que ouvintes fizeram o

julgamento de estímulos no PB e no IN produzidos pelo mesmo falante. Espera-se que os ouvintes identifiquem possíveis mudanças realizadas pelos falantes, e que descrevam as características modificadas, sejam elas vocais, emocionais ou de personalidade.

Por fim, o objetivo geral deste estudo é relacionar os resultados da produção da voz com os da percepção da qualidade vocal. Assim, espera-se que, qualquer mudança em função da língua realizada pelo falante no experimento de produção, seja de alguma forma percebida pelo ouvinte no experimento de percepção.

1.3 PERGUNTAS DE PESQUISA E HIPÓTESES

A fim de investigar a produção da voz de falantes bilíngues de PB e IN e a percepção da qualidade de voz por ouvintes leigos, foram elaboradas perguntas de pesquisa e hipóteses:

P1. Os brasileiros bilíngues produzem diferentes tipos de fonação⁸ nas vogais do IN em relação às do PB?

H1: Não há diferenças entre as medidas espectrais de curto termo produzidas em PB em comparação à produção no IN⁹.

Essa hipótese será testada pelas medidas de declínio espectral ($H1^*-H2^*$, $H2^*-H4^*$, $H4^*-2\text{kHz}$ e $2\text{kHz}-5\text{kHz}$) realizadas na vogal tônica das palavras dissílabas inseridas na frase-veículo, nas duas línguas.

P2: As vozes em inglês dos brasileiros bilíngues apresentam diferenças nas características espectrais em comparação com as vozes em PB? Há diferenças entre as tarefas (leitura de texto e fala semiespontânea) desempenhadas?

H2.1: Haverá diferenças entre as características espectrais da produção em IN em comparação com a produção em PB.

⁸ Medidas de fonação e medidas espectrais de curto termo serão utilizadas como sinônimos neste trabalho.

⁹ Por não haver estudos comparando a fonação alofônica na produção entre línguas por falantes bilíngues, será testada a hipótese nula – H_0 .

Essa hipótese foi baseada em estudos de Bruyninckx *et al.* (1991; 1994), Todaka (1995), Ng *et al.* (2012) e Camargo *et al.* (2013), e será testada pelas medidas de declínio espectral em ELTs obtidas nas amostras de leitura de texto e fala semi-espontânea nas duas línguas.

H2.2: Haverá diferenças nas características espectrais entre a produção em PB e em IN a depender da tarefa desempenhada pelo participante (leitura de texto ou fala semiespontânea).

Essa hipótese foi baseada no estudo de Byrne *et al.* (1994), e será testada por medidas espectrais e de f0 nas amostras de leitura de texto e fala semiespontânea nas duas línguas.

P3: Os brasileiros bilíngues apresentam diferenças de f0 na fala em IN em comparação com a fala em PB? Há diferenças entre as tarefas (leitura de texto e fala semiespontânea) desempenhadas?

H3.1: Haverá diferenças entre as medidas de f0 produzidas em inglês em comparação com as produzidas em PB.

Essa hipótese foi baseada em estudos de Todaka (1995), Altenberg e Ferrand (2006), Ng *et al.* (2012) e Camargo *et al.* (2013), e será testada pela média e extensão da f0 obtidas nas amostras de leitura de texto e fala semi-espontânea nas duas línguas.

H3.2: Haverá diferenças nas medidas de f0 entre a produção em PB e em IN a depender da tarefa desempenhada (leitura de texto ou fala semiespontânea).

Essa hipótese foi baseada em estudos de Bruyninckx *et al.* (1991; 1994) e Harmegnies *et al.* (1991), e será testada por medidas de f0 nas amostras de leitura de texto e fala semiespontânea nas duas línguas.

P4: Haverá diferenças entre as produções dos participantes masculinos em comparação com as participantes femininas em IN e em PB?

H4: As produções de homens e mulheres, tanto nas características espectrais quanto no uso da f0, serão diferentes em IN e em PB.

Essa hipótese foi baseada em estudos de Todaka (1995), Altenberg e Ferrand (2006), Ng *et al.* (2012), e será testada por medidas espectrais e de f0 nas amostras de leitura de texto e fala semi-espontânea nas duas línguas.

P5: A produção da voz em IN dos brasileiros bilíngues está relacionada com as variáveis de idade de aquisição e experiência em países de língua inglesa?

H5: As características vocais da produção em IN estarão relacionadas com as variáveis de idade de aquisição e experiência em países de língua inglesa.

Essa hipótese foi baseada em estudos sobre a aquisição de características prosódicas de L2 como os de Trofimovich e Baker (2006; 2007), Flege (2009) e Huang e Jun (2009), e será testada entre as variáveis relacionadas às características vocais dos bilíngues e as variáveis de idade de aquisição e experiência no exterior.

Com relação ao experimento de percepção, por não haver estudos que investiguem a discriminação de um mesmo falante produzindo línguas distintas, foram apenas levantadas as seguintes perguntas:

P6: Os ouvintes bilíngues leigos percebem diferenças entre as emissões em PB e em IN de um mesmo falante bilíngue?

P7: De que forma os ouvintes leigos quantificam as diferenças entre as emissões em PB e em IN de um mesmo falante bilíngue? Como eles descrevem essas diferenças?

A fim de responder às perguntas de pesquisa e testar as hipóteses propostas, foram desenvolvidos os experimentos de produção e percepção da qualidade de voz de falantes brasileiros bilíngues de PB e IN. Em seguida, os resultados desses dois experimentos foram tratados estatisticamente.

1.4 ORGANIZAÇÃO DA TESE

Esta tese está organizada em sete capítulos. O capítulo 2 apresenta uma visão geral sobre a prosódia, que faz parte da grande área dos estudos fônicos das línguas. Nesse capítulo, são descritas as características prosódicas das línguas, que são geralmente medidas por parâmetros fonético-acústicos de frequência fundamental, intensidade, duração e qualidade de voz. São também tratados nesse capítulo os níveis funcionais das características prosódicas das línguas, a relação entre a qualidade de voz e a pragmática, e aspectos prosódicos paralinguísticos.

O capítulo 3 apresenta uma revisão sobre a produção da voz e a percepção da qualidade de voz, começando pela definição dos termos utilizados na pesquisa sobre o tema. Além disso, são apresentados os métodos mais utilizados para medir a produção da voz e a percepção da qualidade de voz. A fonação é descrita nas suas características fisiológicas e nos diferentes padrões de vibração das pregas. Um modelo psicoacústico para a análise da qualidade de voz é descrito e, em seguida, são apresentadas as características e as medidas acústicas da qualidade de voz adotadas neste trabalho.

O capítulo 4 apresenta uma revisão sobre a aquisição de aspectos prosódicos da L2. Primeiramente, são descritos os fatores que influenciam esse processo, como a idade de aquisição, a experiência linguística, o ambiente de aprendizagem e os fatores sociais e afetivos. Em seguida, são apresentados os estudos sobre qualidade de voz em falantes bilíngues no que concerne a estudos de produção. Por fim, são apresentados aspectos da pesquisa sobre a percepção da qualidade de voz por falantes bilíngues.

O capítulo 5 descreve a metodologia da presente pesquisa em relação ao desenho dos experimentos de produção e percepção da qualidade de voz por falantes bilíngues. Primeiramente, são listadas as hipóteses do estudo. Depois disso, o experimento de produção é detalhado quanto aos seguintes aspectos: os *corpora* utilizados, a entrevista proposta, os participantes, a avaliação do sotaque dos brasileiros por falantes nativos de IN, as gravações, e o procedimento de análise dos dados. A seguir, são descritas as etapas da realização do experimento

de percepção: os ouvintes, os estímulos utilizados, as instruções e as tarefas realizadas pelos participantes.

O capítulo 6 traz os resultados e a discussão do experimento de produção e do experimento de percepção. São apresentados os resultados obtidos por análises estatísticas descritiva e inferencial com *design* intra-sujeito. Além disso, foram realizadas análises estatísticas na comparação entre a produção dos homens e das mulheres, e na correlação entre as variáveis dependentes e as independentes como idade de aquisição, experiência em país de língua inglesa e grau de sotaque atribuído for falantes nativos de IN. Ademais, foram relacionados os dados obtidos com os resultados do experimento de produção, a fim de verificar se uma eventual mudança na produção da voz em IN é percebida pelos ouvintes leigos.

Por fim, o capítulo 7 trará um resumo dos resultados obtidos na pesquisa, e as principais conclusões do estudo. Também, são reportadas as limitações do estudo e sugeridas possíveis pesquisas futuras sobre a produção da voz e a percepção de sua qualidade no que concerne à fala de sujeitos bilíngues.

2 PROSÓDIA

A prosódia constitui uma parte fundamental das interações verbais, já que é por meio dela que veiculamos grande parte das nossas intenções comunicativas e também nossas atitudes e emoções. Ela não somente é responsável por grande parte da variabilidade encontrada na fala, mas também fornece muitas informações concernentes ao significado da frase e à intenção do falante. Como descreve Barbosa (2012),

(...) a prosódia está, no cenário de pesquisa atual, associada a fatores linguísticos como acento, fronteira de constituinte, ênfase, entoação e ritmo, a fatores paralinguísticos como marcadores discursivos (e.g., “né”, “entendo”, “an-han”) e atitudes proposicionais (e.g., “confiante” e “duvidoso”) e sociais (e.g., “hostil” e “solidário”), além de tratar de fatores extralinguísticos como as emoções. Todos esses fatores se combinam com aspectos sociais e biológicos indiciais como gênero, faixa etária, classe social, nível de escolaridade, entre outros. (BARBOSA, 2012, p. 13)

A visão de Barbosa (2012) é atual e abrangente, abrigando uma ampla variedade de funções prosódicas e refletindo a complexidade dessa área da linguística.

Apesar da grande importância da caracterização prosódica das línguas, estudos relacionados à prosódia são escassos em comparação a outras áreas da linguística, já que não contam com um objeto de estudo bem definido, relativamente estável e que se possa quantificar. Segundo Kreiman e Sidtis (2011, p. 261), muito do material prosódico é fluido e contínuo por natureza, o que dificulta o desenvolvimento de descrições estruturais das línguas. Lehiste (1970, p. 1) afirma que “certo grau de imprecisão parece caracterizar a maioria das discussões sobre características prosódicas”¹⁰, e que essa imprecisão se estende à própria definição do que é a prosódia. Para a autora, o termo ‘características prosódicas’ pode ser usado como sinônimo de “características suprasegmentais”. Essa mesma posição é adotada por Cruttenden (1997) quando descreve que “características (prosódicas) geralmente se estendem por enunciados mais longos

¹⁰ ...a certain degree of vagueness seems to characterize most discussions of prosodic features.

do que apenas um som e são, portanto, frequentemente chamados de suprasegmentos”¹¹, e que muitas vezes o termo ‘característica prosódica’ é usado como sinônimo de ‘característica suprasegmental’. Há outros autores que preferem separar as características suprasegmentais das características prosódicas, como é o caso de Cagliari (1992, p. 137). Para ele, elementos suprasegmentais são aqueles que modificam o segmento e são caracterizados como articulações secundárias como a labialização, a palatalização e a nasalização. Já os elementos prosódicos caracterizam unidades maiores que o segmento, como a sílaba e a sentença.

Além desses dois termos, Hirst e Di Cristo (1998) acrescentam um terceiro à discussão, argumentando que o termo ‘entoação’ tem sido frequentemente usado como sinônimo de prosódia e que quando a distinção é feita entre os dois termos, não é clara e explícita. Tais autores propõem que o termo prosódia seja usado no seu senso mais geral possível para englobar tanto os sistemas cognitivos abstratos quanto os parâmetros físicos. Já o termo entoação abarcaria não somente a percepção do *pitch*¹², mas também questões relacionadas à duração e ao ritmo das línguas. Para Barbosa (2012), outra perspectiva seria a de igualmente considerar a prosódia como termo geral, referente aos domínios lexicais e pós-lexicais, mas de conceber a entoação como a variação do *pitch* ao longo do enunciado. Por essa perspectiva, o ritmo está vinculado à prosódia, mas não à entoação, já que aquele depende das variações temporais das sílabas.

Além das diferenças na utilização da nomenclatura, há também diferenças na concepção do que seja prosódia. Há autores, como Shattuck-Hufnagel e Turk (1996), que concebem a prosódia de forma mais abstrata, no que diz respeito à organização fonológica dos segmentos em constituintes de nível mais alto e ao padrão de proeminências relativas dentro desses constituintes. Sob outra perspectiva, Cutler *et al.* (1997) acreditam que a prosódia refere-se à realização de características suprasegmentais físicas da língua, como *pitch*, volume e duração.

¹¹ ...features generally extend over stretches of utterances longer than just one sound and are hence often referred to as suprasegmentals.

¹² O termo *pitch* do IN pode ser traduzido como “altura melódica”, segundo Cordula (2012). Porém, a maioria dos pesquisadores prefere utilizar o termo em IN, como é o que fazemos neste trabalho.

Assim, a estrutura prosódica de um enunciado é produzida por efeitos relacionados à duração, amplitude e frequência do sinal acústico da fala.

Kreiman e Sidtis (2011, p. 261) utilizam o termo prosódia como sinônimo de qualidade de voz¹³. Elas justificam tal escolha argumentando que, hoje em dia, já está bem documentado que as características paralinguísticas da voz realizam funções linguísticas importantes e que, portanto, a separação teórica tradicional entre os aspectos linguísticos e o paralinguísticos da voz não mais se sustenta. Ou seja, a visão de que a informação linguística compreende os níveis fonológico, morfológico, sintático e semântico, e que a informação paralinguística revela atitudes e emoções na fala deu lugar à visão de que características vocais operam em todos os níveis da linguagem. As autoras afirmam que:

O escopo do estudo da linguagem tornou-se muito mais amplo recentemente, expandindo além da fonologia, sintaxe e semântica, para incluir a pragmática da comunicação, que focaliza a fala social e no discurso e examina princípios que governam o uso real da língua. (...) é difícil enfatizar suficientemente o papel da qualidade de voz na pragmática do uso diário da língua (...).¹⁴ (KREIMAN e SIDTIS, 2011, p. 261).

Segundo as autoras, a qualidade de voz é a matéria-prima da prosódia da fala e, hoje em dia, há um grande número de pesquisas sobre esse tema sendo realizadas por diferentes áreas acadêmicas – ciências da fala, psicolinguística e neuropsicologia.

Neste trabalho, adotamos a visão de prosódia de Barbosa (2012), que inclui tanto funções linguísticas como expressivas, considerando-se que a análise prosódica pode ser feita tanto do ponto de vista fonológico quanto fonético. Tal visão está em consonância com o posicionamento de Kreiman e Sidtis (2011), que acrescentam a ideia de que a qualidade de voz é a matéria-prima de todas as funções da prosódia – sejam elas linguísticas ou paralinguísticas. Assim, utilizamos os termos prosódia ou características prosódicas, considerando-os sinônimos, e

¹³ As autoras explicam que a qualidade de voz pode ser concebida de maneira mais ampla ou mais restrita. Tal discussão é aprofundada no capítulo sobre qualidade de voz.

¹⁴ *The scope of language study has also become much broader in recent times, expanding beyond phonology, syntax, and semantics to include the pragmatics of communication, which focuses on social speech and discourse and examines principles governing actual language use. (...) it is difficult to overemphasize the role of voice quality in the pragmatics of language use (...).*

focalizamos suas propriedades físicas no que concerne à qualidade de voz, reconhecendo a indissociável relação entre aspectos linguísticos e paralinguísticos presente no estudo das línguas.

2.1 CARACTERÍSTICAS PROSÓDICAS DAS LÍNGUAS

Kreiman e Sidtis (2011, p. 261) afirmam que:

A prosódia envolve a média do *pitch*, bem como sua variabilidade (ou a média e variabilidade da frequência fundamental), média e variação do volume (ou intensidade), uma grande gama de fatores temporais que determinam a taxa de elocução e o ritmo, e a qualidade de voz (por exemplo, crepitação e sopro, que funcionam sutilmente – e às vezes não tão sutilmente – na fala do dia-a-dia para comunicar o significado)¹⁵. (KREIMAN e SIDTIS, 2011, p. 261).

Esses elementos que compõem as características prosódicas das línguas são, em geral, correlatos acústicos ou perceptuais da produção da fala. A seguir, trataremos de cada um desses componentes sob as perspectivas acústicas e perceptuais.

2.1.1 *Pitch* e frequência fundamental

O *pitch* é considerado um dos componentes mais importantes da caracterização prosódica de uma língua e é comumente referido como correspondente da ‘melodia da fala’. Ocupando uma escala contínua entre ‘alto’ e baixo¹⁶, o *pitch* é o correlato perceptual da frequência fundamental (f0) e que, por sua vez, é uma medida acústica relacionada à taxa de repetição do ciclo de uma onda sonora em um segundo. Cada abertura e fechamento das pregas vocais

¹⁵ *Prosody traditionally encompasses average pitch and pitch variability (or the mean and variability of fundamental frequency), loudness (or intensity) mean and variation, the large array of temporal factors that determine perceived speech rate and rhythm, and voice quality narrowly defined (for example, creakiness and breathiness, which function subtly – and sometimes not so subtly – in everyday speech to communicate meaning).*

¹⁶ Em termos de frequência, e não de volume.

provoca um pico de pressão do ar, criando uma onda sonora. Quanto mais rápido as pregas vocais vibrarem, mais alto é o *pitch*. A taxa de vibração das pregas vocais depende de sua massa e rigidez: quanto maior a massa, mais baixa é a frequência; quanto mais estiradas as pregas estão, mais alta é a frequência. Essa é a razão pela qual os homens têm f_0 mais baixa que mulheres e crianças: média de 115 Hz para eles, 220 Hz para mulheres e 280 Hz para crianças. Contudo, para qualquer falante, a taxa de vibração aumenta quando as pregas estão estiradas e diminui quando estão relaxadas.

Embora a f_0 seja crucial para a percepção do *pitch*, essa correspondência parece ser não-linear. A habilidade que nós, seres humanos, temos de estimar diferenças precisas em movimentos de *pitch* é limitada, pois somente conseguimos identificar características de um som em comparação a outro. A f_0 acontece apenas em sons vozeados, mas mesmo os não-vozeados influenciam a f_0 dos sons vozeados adjacentes. Segundo Cruttenden (1997), as vogais geralmente têm f_0 maior se precedidas de consoantes não-vozeadas do que se precedidas de vozeadas. Além disso, segundo o mesmo autor, o pico da f_0 localiza-se no início da vogal quando esta é precedida por consoantes não-vozeadas, e no meio da vogal quando precedida por consoantes vozeadas. As vogais em si têm o que é chamado de “*pitch* intrínseco”, ou seja, de acordo com Lehiste (1970), vogais altas têm f_0 mais alta. Kreiman e Sidtis (2011, p. 262) asseguram que, embora a frequência fundamental seja quantificável para a maioria dos enunciados, não há formalismos que representem sistematicamente as variações de *pitch* utilizados na fala.

2.1.2 Volume, intensidade e amplitude

O volume de um som, assim como é percebido pelo ouvinte, está relacionado à força expiratória utilizada pelo falante (CRUTTENDEN, 1997). Os correlatos acústicos do volume são a intensidade (medida em energia por uma unidade de área) e a amplitude (medida da extensão da onda sonora). Kreiman e Sidtis (2011) reportam que a variação do volume geralmente acompanha a variação do *pitch*, de maneira que quando este aumenta de frequência, aquele sobe também. Porém, a variação do volume está bastante relacionada com a

distância entre o falante e o ouvinte. De acordo com Cruttenden (1997), a relevância do volume ou intensidade como uma característica prosódica é de difícil avaliação, já que há diversas influências na intensidade absoluta de um som em uma sílaba ou enunciado. Outra dificuldade é que a relação entre intensidade e o volume percebido pelo ouvinte nem sempre é linear. Embora as variações de intensidade sejam importantes para a realização das diferenças entre sílabas tônicas e átonas, não há línguas que utilizem apenas a intensidade como uma pista exclusiva de algum contraste fonológico. Contudo, a intensidade é utilizada para fornecer contrastes emocionais e pragmáticos nas línguas, podendo veicular intenções como relevância, raiva ou tristeza.

2.1.3 Duração e pistas temporais

De maneira geral, a produção e a percepção da fala estão sempre relacionadas ao tempo. Os fatores temporais da prosódia compreendem um conjunto de pistas e funções, tais como a duração de segmentos, sílabas e palavras, taxa de elocução, pausas, padrões acentuais e rítmicos (KREIMAN e SIDTIS, 2011, p. 274). Sob o ponto de vista físico, segundo Lehiste (1970, p. 9), “a fala é constituída de variações nos padrões acústicos em função do tempo”¹⁷. Como aponta essa mesma autora, quando a duração é analisada sob o ponto de vista prosódico, o alcance dessa perspectiva potencialmente será mais abrangente que apenas em um segmento. As medidas temporais mais utilizadas são as durações absoluta e relativa dos segmentos, sílabas e pausas, além da taxa de elocução. Segundo Cruttenden (1997), a relevância da duração enquanto característica prosódica é difícil de avaliar, já que há inúmeros fatores que influenciam a duração absoluta de segmentos e de sílabas. A duração também é fortemente influenciada por fatores como a duração intrínseca das vogais e consoantes, o ambiente fonético e a qualidade das vogais, bem como elementos suprasegmentais.

¹⁷ ...speech consists of variations in the acoustic patterns as a function of time.

2.1.4 Qualidade de voz

A qualidade de voz é a característica prosódica mais sutil de todas, mas não menos essencial na construção dos significados linguísticos. A dificuldade tem sido, entretanto, em avaliar e descrever a qualidade de voz por falta de instrumentos de análise que capturem a interação entre o ouvinte e a fala. Grande parte das pesquisas sobre qualidade de voz tem utilizado métodos auditivos impressionísticos, justamente pela escassez de modelos que permitam descrevê-la. Esses métodos vão desde listas de termos para descrever a qualidade da voz (para mais informações, ver Orlikoff, 1999) a protocolos como o GRBAS (para mais informações, ver Hirano, 1981). O acesso à qualidade de voz tem sido feito via medidas acústicas, principalmente no que diz respeito a medidas de F0, intensidade e duração, mesmo que as relações entre esses parâmetros e as suas funções linguísticas nem sempre sejam claras e diretas. Segundo Kreiman e Sidtis (2011), a maioria das pesquisas em linguística sobre qualidade de voz diz respeito aos tipos de fonação empregados pelas línguas em nível fonológico, embora haja um crescente interesse pelo uso de diferentes qualidades de voz em outros níveis linguísticos. Como exemplo, as autoras ressaltam que o uso da voz crepitante geralmente indica fronteiras prosódicas (como o final de sentença) e mudança de tópico. Contudo, diferentemente de outros aspectos prosódicos, as qualidades de voz não se beneficiam de um sistema notacional para utilizar na avaliação impressionística do sinal vocal. A definição e a mensuração da qualidade de voz serão tratadas no capítulo 3.

2.2 NÍVEIS FUNCIONAIS DAS PISTAS PROSÓDICAS

As características prosódicas das línguas apresentam diferentes funções comunicativas que podem operar como pistas discretas, gradientes ou idiossincráticas. Kreiman e Sidtis (2011) descrevem essas características, apontando os domínios sobre os quais operam, estabelecendo seus níveis funcionais e seus *status* (FIGURA 1):

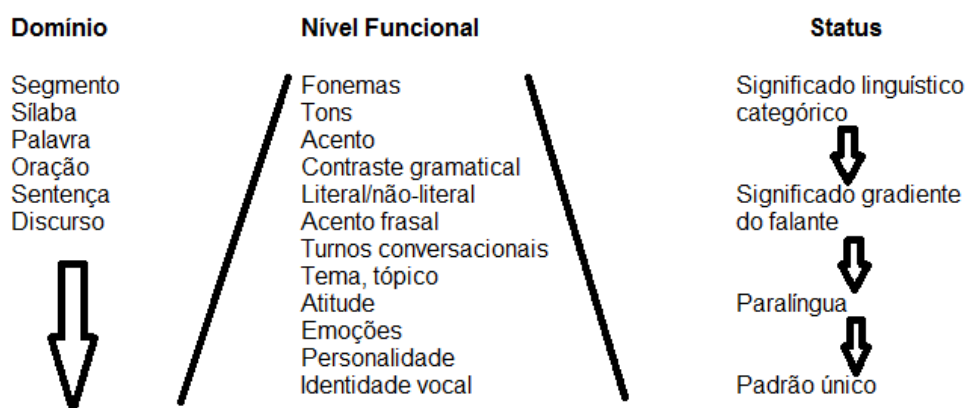


Figura 1 - A PIRÂMIDE DA PROSÓDIA
 FONTE: Adaptado de Kreiman e Sidtis (2011, p. 277)

Nessa representação, as pistas prosódicas da fala são divididas por unidade linguística (domínio), função linguística (nível), propriedade primária (discreta ou contínua) e tipo de significado (*status*). No topo dessa pirâmide estão os contrastes linguísticos fornecidos pela qualidade de voz, duração, amplitude e *pitch*. Tal domínio inclui contrastes segmentais baseados em tipos de fonação, tons em línguas tonais, entoação de perguntas *versus* declarativas, entre outros. Mais abaixo, a partir do acento frasal até a identidade vocal, fica mais difícil de identificar ou descrever quais pistas servem para cada função prosódica. Isso acontece, pois tais pistas são gradientes e por vezes se sobrepõem; ou seja, muitas vezes as pistas são intercambiáveis e podem indicar mudanças de tópico ou unidades discursivas, e os atributos prosódicos associados a essas funções variam idiossincraticamente de indivíduo para indivíduo, e enunciado para enunciado. Além disso, tais funções refletem aspectos do significado que o falante pretende fornecer, bem como o estilo comunicativo e as características pessoais. Por fim, na base da pirâmide encontram-se as características paralinguísticas como a atitude, a emoção e a personalidade, que são veiculadas por uma variedade de características prosódicas, tornando difícil o estabelecimento de correlações consistentes entre essas funções, o significado que elas trazem e a estrutura linguística das línguas.

Barbosa (2012, p. 17) propõe que as características prosódicas das línguas têm três funções distintas: a) funções discursivas dialógicas (e.g. troca de turno) e

não-dialógicas (e.g. modalidades, como o monólogo); b) funções demarcativas, que assinalam fronteiras prosódicas; c) funções de marcação de proeminência (e.g. saliência de um constituinte sobre o outro). Por outro lado, no plano expressivo, o autor aponta funções prosódicas distintas: a) as atitudinais (e.g. postura interpessoal); b) as afetivas (e.g. tristeza, humor); c) as indiciais (e.g. gênero e origem social).

No plano linguístico, como exemplo da função discursiva, Barbosa (2012) ilustra a diferença entre um enunciado assertivo e outro interrogativo, descrevendo-a pela subida da f_0 no final do enunciado na frase interrogativa. Já a função demarcativa pode ser exemplificada pelo uso da pausa (ou pelo preenchimento dela) para assinalar fronteiras prosódicas. A função de proeminência pode ser ilustrada pela ênfase que damos a uma informação no enunciado para torná-la foco da mensagem, aumentando sua energia de produção pela maior duração e subida de *pitch*. No plano expressivo, o autor afirma que as características prosódicas atitudinais, afetivas e indiciais são verificadas em todo o enunciado, já que esses parâmetros são dificilmente disfarçáveis na produção da fala.

Por conta da complexidade do sinal da fala, que envolve o significado linguístico, a intenção do falante, questões paralinguísticas e de identidade, tentar separar essas funções nas características prosódicas das línguas pode ser uma tarefa intimidante. Mesmo assim, falantes e ouvintes detêm uma capacidade incrível em decifrar o significado das pistas prosódicas contidas em cada enunciado.

2.3 QUALIDADE DE VOZ E A PRAGMÁTICA

A pragmática é o ramo da linguística que estuda o uso da linguagem em situações de comunicação. A pesquisa nessa área se concentra em descrever como os falantes cooperam de diferentes maneiras para alcançar a comunicação verbal de maneira satisfatória, fazendo o uso de suposições e inferências. É por meio da qualidade de voz que diversos recursos pragmáticos são sinalizados nas

situações de comunicação, como por exemplo: trocas de turno e mudança de tópico; autenticidade e sinceridade; confiança e polidez; deferência e assertividade.

Segundo Kreiman e Sidtis (2011), há uma série de marcas discursivas que são manifestadas pela qualidade de voz. O tópico e o tema são moldados pela qualidade vocal, assim como o acento frasal para indicar quando uma informação é nova ou dada. Modulamos nossa voz pelo aumento da duração, do *pitch* e/ou da intensidade quando apresentamos uma nova informação no discurso. Mudanças no acento frasal também acarretam diferenças entre um pedido e uma ordem, já que a mudança na entoação altera a força ilocucionária do enunciado. Contudo, as pistas vocais utilizadas como recurso pragmático são difíceis de classificar e de codificar, pois muitas dessas pistas são também utilizadas para fornecer emoções e atitudes na fala. Mesmo assim, como parte da nossa competência pragmática, temos a capacidade de discernir os diferentes níveis de ênfase e importância das pistas prosódicas em situações de comunicação.

Outro papel importante da prosódia é o de sinalizar unidades maiores do discurso, como por exemplo, o parágrafo. Essa unidade, quando lemos um texto ou contamos uma história, apresenta estrutura bem definida, sendo parte dessa estrutura veiculada pela modulação da prosódia. Geralmente, falamos com intensidade e velocidade maiores quando iniciamos um novo tópico e utilizamos *pitch* mais baixo no fim do enunciado; essas mudanças fornecem pistas robustas para os ouvintes (KREIMAN, 1982, p.172). A estrutura das unidades do discurso proporcionadas pelas pistas prosódicas é crucial para a compreensão, já que um dos papéis mais importantes da prosódia é o de auxiliar na codificação semântica e na retenção de itens lexicais na memória (COHEN *et al.*, 2001).

2.4 PROSÓDIA E ASPECTOS PARALINGÜÍSTICOS

Segundo Karpinski (2012, p. 38), embora o poder de descrição e de flexibilidade da análise linguística tenha aumentado significativamente nas últimas décadas, parece ainda haver uma fronteira entre o que pertence ao domínio linguístico e ao que é considerado “acessório” para a comunicação. Mesmo assim,

há um crescente interesse na pesquisa em torno do uso da linguagem e interação, pois muito embora muitas das propriedades salientes da fala não necessariamente correspondam a unidades previstas na representação do componente linguístico das línguas, são pistas que fornecem informações sobre o falante, suas características psicossociais e atitudinais, tanto em relação ao tópico, quanto ao interlocutor. O interesse pelo estudo da fala, variação fonética e aspectos sociais e psicológicos da comunicação cresceu ainda mais, a partir da década de 80, por conta dos avanços tecnológicos que possibilitaram o desenvolvimento da síntese e reconhecimento de fala (KARPINSKI, 2012, p. 39).

Para Crystal (1975, p. 162), a paralingua é vista como um elo entre formas de comportamento não-linguístico e os diversos níveis que se costuma estudar na linguística. De acordo com esse autor, acima do nível fonológico da fala, há certos efeitos vocais que têm um papel crucial para a comunicação, e são geralmente classificados como características paralinguísticas. Em sua concepção de paralingua, Crystal (1975) não inclui os aspectos relacionados à comunicação visual (imprescindíveis para a comunicação), justificando que os fatores vocais são, por si só, complexos o suficiente. Essa complexidade se refere à função da paralingua na comunicação, já que ela fornece informações gramaticais, atitudinais e sociais. Defendendo a importância dos estudos paralinguísticos, Crystal afirma:

Certamente, observações das reações das pessoas em relação à linguagem no seu uso diário sugerem que os fenômenos paralinguísticos, longe de serem marginais, são frequentemente os determinantes primários do comportamento em uma interação, por vezes atribuindo um papel secundário aos chamados aspectos 'conotativo' ou 'denotativo' do uso da língua.¹⁸ (CRYSTAL, 1975, p. 164)

Ainda para Crystal (1975), uma das principais funções da paralingua é a de expressar emoção. Porém, esse autor considera que o uso de características paralinguísticas como marcadores da estrutura gramatical do discurso tem uma função ainda mais importante. Assim como a pontuação é uma convenção crucial para a estruturação da comunicação escrita, garantindo certo nível de

¹⁸ *Certainly, observations of people's everyday reactions to language suggest that paralinguistic phenomena, far from being marginal, are frequently the primary determinants of behaviour in an interaction, sometimes pushing the so-called 'cognitive' or 'denotative' aspects of the utterances used into a secondary role.*

inteligibilidade, a modulação de fatores paralinguísticos organiza a cadeia da fala em unidades que viabilizam a comunicação oral. Outra função paralinguística é a de sinalizar a intenção do falante, como por exemplo, persuadir, irritar ou alegrar. Além disso, a paralingua fornece meios para expressarmos estados psicológicos, como dominância, liderança e submissão.

Como relata Karpinski (2012, p. 39), linguistas têm consciência da questão problemática que é estabelecer fronteiras entre o que é linguístico e o que é paralinguístico. No processo da fala, sinalizamos uma série de significados linguísticos por mudanças de *pitch*, duração e intensidade. Porém, quando um falante escolhe enfatizar alguma informação do discurso pelas mudanças de *pitch*, duração ou intensidade, estão também implicadas em sua fala determinadas atitudes, ou seja, informação paralinguística. Dessa forma, como apontam Kreiman e Sidtis (2011, p. 294), a informação paralinguística de característica atitudinal origina-se de diversos parâmetros prosódicos de caráter linguístico que, por sua vez, desempenham múltiplas funções na organização linguística. Portanto, se não se consegue atribuir inequivocadamente a mudança dessas dimensões prosódicas a fatores linguísticos ou aos paralinguísticos de maneira exclusiva, não se pode pensar em uma divisão entre esses dois domínios. Consequentemente, a qualidade de voz como concebida por Kreiman e Sidtis (2013), atua como matéria-prima da prosódia, e tem como papel veicular tanto propriedades linguísticas como paralinguísticas.

Muitas das propriedades prosódicas que usamos na fala são entendidas como a expressão de atitudes e emoções que variam de língua para língua e de cultura para cultura. Kreiman e Sidtis (2011, p. 296) afirmam que enunciados desprovidos de propriedades atitudinais mal existem. Segundo elas, as atitudes refletem o ponto-de-vista do falante no enunciado e pertencem tanto ao domínio linguístico quanto ao paralinguístico. Autenticidade, sarcasmo, reprovação, sinceridade são algumas das atitudes fornecidas pelos falantes e percebidas por seus interlocutores. Por exemplo, a resposta à pergunta “Onde está [substantivo]?” (“Está no/na [local]”) pode soar informativa, desafiadora, defensiva, entre outras atitudes, a depender da modulação do *pitch*, duração intensidade e qualidade de voz empregada pelo falante. O valor que se atribui às diversas características

atitudinais veiculadas pela qualidade vocal reflete normas e tendências sociais, e também o preconceito. Isso prova que a informação atitudinal proveniente das propriedades vocais do falante tem um forte efeito na percepção do significado linguístico. Da mesma forma, Kreiman e Sidtis (2011) relatam que informações sobre as emoções do falante são raramente veiculadas por meio da fonologia, morfologia ou sintaxe; tais características são fornecidas quase que exclusivamente pela semântica, pela escolha lexical e, principalmente, pela qualidade de voz. Para expressar tristeza, medo, raiva, alegria e tédio, os falantes geralmente modificam o *pitch* (tanto a média, como a extensão e a variabilidade de f0), a intensidade e a taxa de elocução em relação à sua fala natural.

Contudo, apesar do crescente interesse pela pesquisa sobre a relação entre a qualidade de voz e a emoção e atitude, há ainda um conhecimento limitado de como os ouvintes as interpretam na interlocução com o falante. Como resumem Kreiman e Sidtis (2011):

A informação emocional nas interações de fala é disposta em camadas, é fluida e vaga; a percepção dos ouvintes fornece muitos tipos de filtros psicológicos e experienciais; as pistas acústicas são normalmente embutidas em outras pistas físicas, linguísticas e sócio-contextuais. (...) E claro, fatores culturais provavelmente têm um papel significativo no estabelecimento do processo que permite a produção e percepção da emoção (...).¹⁹ (KREIMAN e SIDTIS, 2011, p. 329).

Além das características atitudinais e emocionais, a paralingua também comporta as características de identidade do falante. Laver e Trudgill (1979) propõem que as características indiciais do falante sejam agrupadas em três categorias. São elas:

- (a) as que marcam as características sociais, como afiliação regional; status social e educacional, profissão e papel social;
- (b) as que marcam características físicas, como idade, sexo, e estado de saúde;
- (c) as que marcam as características psicológicas, como a personalidade e o estado emocional. (LAVER e TRUDGILL, 1979, p. 3).

¹⁹ *Emotional information in conversational speech is layered, fluid and vague; listeners' perceptions provide many kinds of psychological and experiential filters; acoustic cues are usually embedded in other physical, linguistic and socio-contextual cues. (...) And, of course, cultural factors are likely to play a significant role in the establishment of processes that allow for production and perception of emotion (...).*

Primeiramente, como soamos depende fortemente das nossas origens geográficas, do fato de sermos estudados ou não, ou de termos determinado papel social. Como soamos também depende se somos homens ou mulheres, da nossa orientação sexual, e dos nossos atributos físicos. Além disso, como soamos depende, sobretudo, de nossa personalidade, do nosso estado de espírito e do nosso humor. Por conta dessa relação entre a fala e as características físicas e psicossociais, o ouvinte percebe a qualidade de voz como uma pista para julgar todas essas características indiciais que fornecemos quando falamos. Somos capazes de mudar a nossa voz para comunicar diferentes intenções comunicativas e o fazemos com relativa facilidade. Essa mudança pode ser resultante da alteração do fluxo de ar que vem dos pulmões, da forma como as pregas vibram e da variação do tamanho e do formato do trato vocal, ações essas que podem ocorrer individualmente ou em conjunto. Um número aparentemente pequeno de alterações possibilita uma enorme gama de características vocais que podemos produzir. É importante ressaltar, contudo, que essas mudanças são condicionadas pela fisiologia de cada trato vocal; a extensão da variação de volume, *pitch* e qualidade são determinados pela anatomia vocal inerente do falante.

Assim, a literatura convencionou discernir as características indiciais da qualidade de voz em duas partes: as que são “orgânicas” e as que são “adquiridas”. As primeiras são aquelas resultantes da configuração anatômica do falante, e que, conseqüentemente, são relativamente imutáveis, como média de f_0 e frequência de formantes. Por serem motivadas fisiologicamente, as características orgânicas da voz tendem a ser estáveis para falantes adultos, fornecendo bons índices de identidade, sexo e idade (KREIMAN e SIDTIS, 2011, p. 65). As últimas são padrões habituais que o falante aprende com a experiência linguística, que incluem o sotaque, taxa de elocução, curvas entoacionais, f_0 habitual e qualidade de voz. Essas características são cruciais para sinalizar o vínculo a determinada comunidade linguística e para fornecer atributos pessoais. Portanto, a voz é o produto da composição fisiológica e da origem sociolinguística do falante.

2.5 CONCLUSÃO DO CAPÍTULO

A prosódia é parte inerente à fala e imprescindível na comunicação humana. Segundo Karpinski (2012, p. 40), “todos os enunciados, até mesmo os gerados por computador, têm um determinado ritmo (mesmo que monótono), entoação (mesmo que *flat*) ou qualidade de voz (mesmo que pouco expressiva) (...)”²⁰. A qualidade de voz permeia todo o sinal da fala e sinaliza diversas funções linguísticas – fonológicas, gramaticais e pragmáticas. Os contrastes fonêmicos ou os tonais são realizados pelo uso fonético da voz, assim como o acento lexical e outros padrões acentuais. Além de diversas funções linguísticas, a prosódia fornece também propriedades paralinguísticas na cadeia da fala. A paralingua é responsável por veicular informações relacionadas às características físicas e psicossociais do falante, bem como suas atitudes e emoções na produção do enunciado. Segundo Pike (1945), as atitudes e intenções veiculadas na comunicação oral são pelo menos tão importantes quanto as informações linguísticas e, às vezes, têm mais importância que estas. O grande desafio parece ser separar na prosódia o que é linguístico e o que é chamado de paralinguístico – já que esses domínios estão fortemente entrelaçados e muitas vezes utilizam-se das mesmas medidas para descrever os dois níveis. Linguistas, na sua maioria, consideram que as propriedades linguísticas sejam discretas, enquanto que as paralinguísticas, gradientes.

Portanto, o presente estudo se situa na descrição de uma característica prosódica – a qualidade de voz – a partir das suas funções paralinguísticas na produção e percepção da fala de bilíngues. Para que se possa investigar esse fenômeno, é necessário entender como as pesquisas definem a qualidade de voz e como se pode medi-la. Assim, o próximo capítulo apresenta possíveis definições e medidas mais utilizadas para analisar a produção da voz e a percepção da sua qualidade.

²⁰ *All utterances, even computer-generated ones, have a particular rhythm (even if monotonous), intonation (even if flat), or voice quality (even if not very striking)...*

3 QUALIDADE DE VOZ

Quando falamos, nossa voz fornece uma série de informações sobre nossas características tanto físicas quanto psicológicas. Podemos reconhecer quando uma pessoa está cansada ou entediada, distraída ou doente, quando quer contar um segredo ou está entusiasmada. Embora nós, como ouvintes, percebamos todos esses elementos que as vozes nos revelam, nosso julgamento nem sempre é preciso; ainda assim, geralmente sabemos se uma pessoa é um homem ou uma mulher, se está cansada ou irritada, doente ou animada, além de percebermos atitudes do falante, como ironia e sarcasmo, competência e credibilidade. A qualidade de voz também nos fornece informações sobre trocas de turno e ajuda a desfazer ambiguidades nos atos de fala. Kreiman e Sidtis (2011, p. 2) apresentam alguns julgamentos que ouvintes fazem sobre as vozes (QUADRO 1):

Características físicas: idade; aparência; condição dental, oral e nasal; condição de saúde, fadiga; identidade; intoxicação; raça, etnia; sexo, orientação sexual; fumante, não-fumante.
Características psicológicas: estímulo (relaxado, apressado); competência; estado emocional, ânimo; inteligência; personalidade; estado psiquiátrico; estresse; veracidade.
Características sociais: escolaridade; ocupação; origem regional; papel na situação da fala; condição social.

QUADRO 1: TIPOS DE JULGAMENTO DA VOZ

FONTE: Kreiman e Sidtis (2011, p. 2), tradução da autora da tese.

Além de todas essas características, certas patologias podem também ser detectadas pela qualidade de voz. Pacientes com alguma alteração geralmente reclamam que sua voz não corresponde à sua identidade, e muitas vezes até evitam falar, resultando em dificuldades de relacionamento (KREIMAN e SIDTIS, 2011 p. 3).

Assim, pela grande quantidade de informações que fornecemos pela voz, há também um grande número de pesquisadores interessados em estudar a produção

e a percepção da voz. Linguística, biologia, ciência da computação, ciência forense, medicina, música, psicologia, ciências da fala e fonoaudiologia são alguns exemplos de áreas que têm desenvolvido pesquisas sobre a voz. Esse interesse de tantas áreas distintas do conhecimento revela a grande importância da voz para a vida humana.

3.1 DEFINIÇÃO DE QUALIDADE DE VOZ

Segundo Kreiman e Sidtis (2011), definir os termos “voz” e “qualidade de voz” não é uma tarefa fácil, já que os dois têm sido frequentemente usados de maneira intercambiável. Essas autoras defendem definições baseadas no uso científico dos termos, de modo que a ‘voz’ tenha uma base física e fisiológica que remete ao sinal acústico, e a ‘qualidade de voz’ seja referente à impressão perceptiva resultante do sinal acústico, de maneira análoga à distinção entre frequência (propriedade física da vibração) e *pitch* (sensação do ouvinte). Essa definição de ‘voz’ ainda permite duas outras possibilidades. A primeira é a de que ‘voz’ seja sinônimo de fonação; ou seja, da vibração produzida pelas pregas vocais. Tal definição, contudo, exclui outras ressonâncias produzidas no trato vocal, como por exemplo, o ruído fricativo. A outra definição refere-se ao uso do termo ‘voz’ como sinônimo de ‘fala’. Essa segunda definição parece, mesmo que de modo geral, incluir o resultado acústico da ação coordenada entre o sistema respiratório e os outros elementos do trato vocal envolvidos na fala. Assim como ‘voz’, o termo ‘qualidade de voz’ pode ser definido de maneira mais específica ou mais abrangente: como o resultado percebido do processo de fonação, ou como a resposta de um ouvinte ao som da fala, respectivamente. A discussão sobre a definição dos termos não para por aí, como apontam as autoras, já que definições dependem do propósito e da perspectiva adotadas pela área específica que as estuda.

Alguns autores optaram por definir qualidade de voz em termos fisiológicos, tanto pela maneira mais abrangente quanto pela específica. Abercrombie (1967, p. 91) define a qualidade de voz como "as características presentes mais ou menos

todo o tempo que uma pessoa está falando: é uma qualidade quase permanente que permeia todo o som emitido da boca"²¹. Para Laver, a qualidade de voz é:

Uma abstração cumulativa de um falante em um período de tempo, caracterizando-se como qualidade, que é acumulada a partir das flutuações momentâneas e espasmódicas de articulações de curto-termo, usadas pelo falante para a comunicação linguística e paralinguística.²² (Laver 1980, p. 1)

Abercrombie e Laver utilizam a definição mais abrangente, que corresponde à percepção do resultado da ação coordenada do sistema respiratório, pregas vocais, língua, mandíbula, lábios e palato mole (KREIMAN *et al.*, 2005). Tais definições, contudo, não especificam a contribuição do ouvinte na interpretação da qualidade de voz que é, por sua vez, um fenômeno perceptual por natureza. Segundo Kreiman *et al.* (2005), diferentes aspectos da percepção da qualidade de voz dependem do contexto, da atenção, da experiência do ouvinte, sendo também afetados pelo tipo de tarefa perceptiva que se desempenhe. Por isso, os julgamentos perceptivo-auditivos de uma voz geralmente não são constantes entre ouvintes distintos.

Outro problema seria a definição da qualidade de voz somente pelo *pitch* e intensidade, medidas unidimensionais que vão de baixo para alto e de forte para fraco. Nossa audição responde a medidas de *pitch* e intensidade de maneira bastante estável, segundo Kreiman e Sidtis (2011), inclusive entre indivíduos. Contudo, qualidade de voz é multidimensional, já que há para ela mais de uma pista perceptiva e variações entre ouvintes podem existir, resultando na impossibilidade de determinação de um conjunto de medidas acústicas fixas. É por isso que a qualidade de voz pode ser considerada como uma interação entre um ouvinte e um sinal, já que o ouvinte capta diferentes pistas acústicas disponíveis no momento para atingir um objetivo perceptivo (KREIMAN *et al.*, 2005). Os ouvintes se utilizam de diferentes estratégias na percepção da qualidade, podendo basear-se em características articulatórias e/ou na familiaridade com a voz.

²¹ *...those characteristics which are present more or less all the time that a person is talking: it is a quasi-permanent quality running through all the sound that issues from his mouth.*

²² *...a cumulative abstraction over a period of time of a speaker characterizing quality, which is gathered from the momentary and spasmodic fluctuations of short-term articulations used by the speaker for linguistic and paralinguistic communication.*

Neste trabalho, adotaremos a definição de “voz” como sinônimo de “fala” para contemplar tanto os aspectos laríngeos quanto os supralaríngeos da produção, e utilizaremos o termo “qualidade de voz” como a resposta de um ouvinte ao som da fala. A escolha por tais termos se justifica pela possibilidade de contemplar tanto os aspectos da produção quanto da percepção dos sons da fala.

3.2 COMO MEDIR A VOZ E SUA QUALIDADE

Além de ser difícil de definir, a qualidade de voz também é difícil de medir, dada a sua natureza perceptiva: "É uma impressão psicológica criada por um estímulo físico, e, portanto, depende tanto da voz como do ouvinte (...)"²³ (KREIMAN *et al.*, 2005 , p. 345). Pelo fato de esta impressão psicológica ser difícil de medir, a pesquisa na área centrou-se no desenvolvimento de listas de rótulos para descrever as vozes, pela presença ou ausência de qualidades específicas. Muitas dessas listas foram criadas para resolver o problema da falta de um modelo padronizado para a descrição da qualidade de voz, mas a maior parte delas tem mais ou menos as mesmas características. Contudo, tais listas podem apresentar níveis diferentes de descrição, misturando impressões visuais, sinestésicas, físicas e estéticas, entre outras. Um dos problemas apresentados pelas listas é que elas são geralmente redundantes e ambíguas, características que levaram os pesquisadores a aplicar a análise fatorial, um procedimento estatístico para reduzir longas listas de características sobrepostas em conjuntos menores e não redundantes. A análise fatorial é considerada mais econômica e mais eficaz do que listas de termos, porque os fatores são independentes um do outro e capturam grande parte da informação nas avaliações escalares. No entanto, esse método de análise pode não ser totalmente adequado, uma vez que depende das escalas e dos estímulos, limitando o alcance da generalização dos resultados. Além disso, muitos estudos que se utilizaram da análise fatorial tinham um número reduzido de estímulos e falantes, sendo comum o julgamento a partir de vogais sustentadas, resultando em variações bastante grandes entre um estudo e outro. Assim, como

²³ *it is the psychological impression created by a physical stimulus, and thus depends on both the listener and the voice ...*

apontam Kreiman e Sidtis (2013), um modelo perceptivo adequado não pode ser estruturado apenas em razão dos estímulos, mas também deve dar conta separadamente das diferenças entre os ouvintes.

Uma resposta à inconstância dos julgamentos a partir de listas de termos e análise fatorial foi a utilização do escalonamento multidimensional (*Multidimensional Scaling* - MDS). Segundo Kreiman *et al.* (2005, p. 350), o MDS permite que o ouvinte avalie a semelhança entre diferentes estímulos de forma direta, produzindo um espaço perceptivo n-dimensional de modo que a distância entre as vozes são proporcionais às semelhanças entre elas (quanto mais similares, mais próximas). As dimensões utilizadas variam desde medidas acústicas (medidas de f0, formantes, intensidade, duração), características físicas do falante (sexo, idade), características articulatórias (nasalidade, rouquidão, soproidade) e até psicológicas (ânimo e masculinidade). O caráter exploratório do MDS traz a possibilidade de relacionar a qualidade de voz como um todo com qualidades específicas da voz. Existem, no entanto, limitações na utilização do MDS, uma vez que não leva em conta a invariância apresentada pelos estímulos. Isso geralmente acontece por causa do pequeno número de estímulos utilizados em estudos. Além desse fato, por falta de evidências empíricas, não fica claro quais dimensões incluir ou excluir em um modelo para a descrição de qualidade vocal.

Como reação a essas limitações, o *Voice Profile Analysis Protocol* (VPAS), desenvolvido por Laver *et al.* (1991), Laver (2000) e Mackenzie-Beck (2005), propôs um modelo para a análise das qualidades de voz que incluiu suas configurações fonética e fisiológica. Esse protocolo é baseado em uma descrição de qualidade de voz foneticamente fundamentada em que o “ajuste”²⁴ atua como unidade analítica. O julgamento perceptivo-auditivo da qualidade de voz é feito com base nas variações de um grupo de ajustes usado como referência – o chamado ajuste neutro. O VPAS é considerado um instrumento de medida sensível à qualidade vocal já que seus parâmetros analíticos dão informações sobre a contribuição de todo o trato vocal na produção da fala (MACKENZIE-BECK, 2005),

²⁴ O ajuste pode ser definido como "uma tendência de longo prazo muscular", relacionado às dimensões fonatória (modos de vibração), supralaríngea (articulatório) e de tensão (laringe e do trato vocal). (CAMARGO e MADUREIRA, 2008)

e não apenas sobre os tipos de fonação. Mackenzie-Beck (2005) observa que o grau de treinamento do juiz é um fator indispensável para criar confiança e, consequentemente, aumentar a confiabilidade do julgamento perceptivo-auditivo.

O maior problema de todos os modelos aqui apresentados é que todos implicam que a qualidade de voz possa ser representada somente por listas de descritores ou dimensões, que haja uma lista de atributos aos quais os ouvintes prestam atenção, e que o mesmo conjunto de parâmetros sirva para todas as vozes (KREIMAN e SIDTIS, 2011, p. 20). Segundo as autoras, um modelo bem definido e embasado teoricamente ainda não foi desenvolvido e bem estabelecido, apesar de anos de pesquisa na área, principalmente porque parece que os ouvintes exploram o sinal acústico cada um de sua maneira. Outra dificuldade é a confiabilidade do julgamento dos ouvintes já que, na sua maioria, não há completa concordância entre julgamentos de uma única voz. Os modelos até então desenvolvidos não preveem o comportamento do ouvinte e não dão conta de fatores como atenção, experiência e respostas enviesadas.

A partir de todas essas dificuldades, pesquisadores da área têm se voltado para outros meios de medir a voz e a qualidade vocal. A alternativa às medidas perceptuais tem sido as medidas instrumentais, como por exemplo: medidas acústicas, medidas aerodinâmicas, eletroglotografia, ressonância magnética, palatografia, ultrassom, e videoestroboscopia. Ao contrário dos protocolos perceptuais, as medidas instrumentais fornecem precisão, confiabilidade e replicabilidade. Mesmo contando com medidas precisas das características da voz, Kreiman *et al.* (2004, p. C166) afirmam que o "desenvolvimento de protocolos instrumentais para medir a qualidade, em última análise, depende da nossa capacidade em definir a qualidade de voz de maneira que fatores cognitivos introduzam a variabilidade das medidas"²⁵. Eles argumentam que, até então, nenhum modelo existente era tão abrangente a ponto de dar conta de maneira precisa das características fisiológicas e acústicas de qualidade vocal por meio da relação entre a análise instrumental e diferenças perceptuais.

²⁵ ...development of instrumental protocols for measuring quality ultimately depends on our ability to define quality in a way that accounts for perceptual factors that introduce variability in listeners' judgements.

3.3 ANATOMIA E FUNCIONAMENTO DAS PREGAS VOCAIS

Uma vez que o ar proveniente dos pulmões flui através da laringe, ele pode ser convertido em som pela vibração das pregas vocais. Estas oscilam, abrindo e fechando, ora bloqueando a passagem do ar e criando assim diferenças de pressão, que ouvimos como som. O processo no qual as pregas vocais modulam o fluxo do ar para a produção de um som é chamado de fonação. Nos sons vozeados, há a vibração das pregas, enquanto que nos sons não-vozeados, as mesmas permanecem afastadas, permitindo o fluxo mais livre do ar pela laringe.

As pregas vocais estão localizadas na laringe, que está suspensa no pescoço por músculos, ligamentos e membranas ao osso hióide. Segundo Kreiman e Sidtis (2011), embora a laringe tenha evoluído para produzir os sons da fala, sua função primária e mais importante é a de impedir a entrada de objetos e fluidos nas vias aéreas e nos pulmões, principalmente durante a deglutição. A laringe abrange um conjunto de cartilagens interconectadas e está localizada abaixo da faringe e acima da traquéia. A maior cartilagem da laringe é a cartilagem tireóideia, que tem formato de escudo na frente e é aberta na parte de trás. A segunda maior cartilagem é chamada cricóideia, que tem o formato de um anel de sinete, com a parte maior voltada para trás. Há ainda as cartilagens aritenóideas, que têm formato de pirâmides de três lados e estão localizadas em cima da parte traseira da cartilagem cricóideia. Segundo Pinho et al. (2014):

(...) o movimento das cartilagens aritenóideas em relação à cartilagem cricóideia se dá por deslizamento nas direções anterior e inferior, ou nas direções posterior e superior. (...) Além disso, quando estas se movem anterior ou inferiormente, ocorre uma outra mudança de direção. As cartilagens aritenóideas, junto com as pregas vocais, movem-se medialmente, no encontro uma da outra (adução). A abdução acontece no sentido contrário (aritenóides movendo-se posterior e superiormente). PINHO et al, 2014, p. 16.

As pregas vocais atravessam a passagem do ar logo abaixo do encaixe entre as cartilagens tireóideia e aritenóideia. As pregas são pequenas: em homens, variam de 17 a 24 milímetros, e em mulheres, de 13 a 17 milímetros; contudo, são muito flexíveis e podem se estender por mais três a quatro milímetros. O espaço

entre as pregas vocais é chamado glote; o espaço acima delas é chamado supraglotal, e o abaixo, subglotal. Logo acima das pregas vocais estão as pregas ventriculares, também chamadas de falsas pregas vocais.

As pregas vocais têm uma estrutura complexa e em camadas, cuja composição é feita de diferentes tecidos, possibilitando que o corpo das pregas seja estirado e enrijecido enquanto que a cobertura que as envolve esteja solta e possa se mover livremente. Os músculos laríngeos realizam a abertura da glote para a respiração ou seu completo fechamento para bloquear as vias aéreas, posicionam as pregas vocais para a produção de sons, ou estreitam a abertura glotal para produzir o sussurro/cochicho. Tais músculos são divididos em adutores (fecham as pregas vocais para realizar a fonação ou bloquear as vias aéreas) e abdutores (abrem as pregas para a respiração). O músculo laríngео mais importante é o tireoaritenóideo (FIGURA 2), que é utilizado para controlar a frequência do som e a mudança na qualidade de voz.

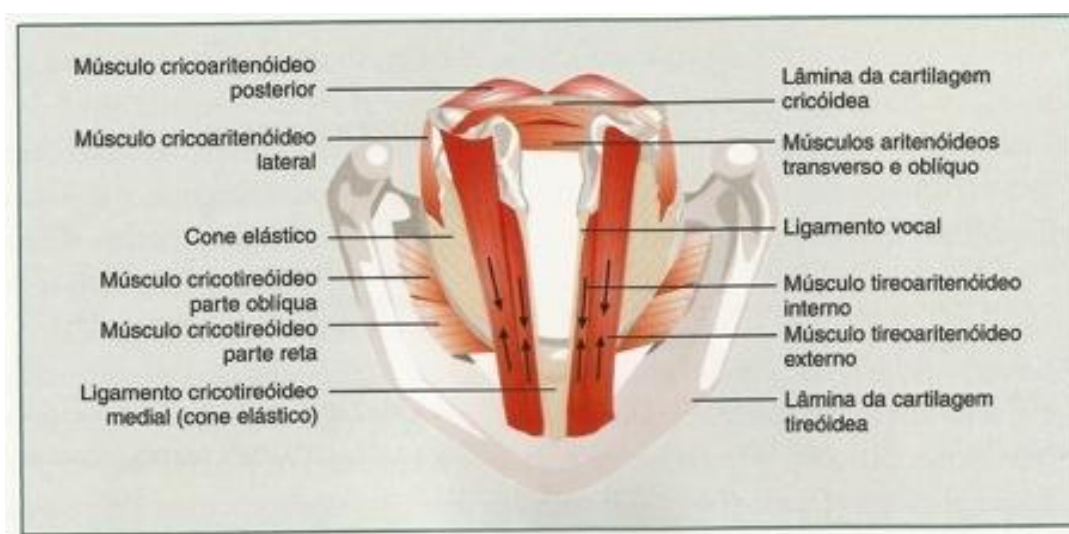


FIGURA 2 - MÚSCULOS E LIGAMENTOS ENVOLVIDOS NA FONAÇÃO
FONTE: PINHO *et al.* (2014, p.16)

Embora os músculos laríngeos tenham um papel imprescindível na fonação, eles não são os únicos responsáveis pela produção dos sons da fala. Para que um som seja produzido, é necessário que haja a combinação da elasticidade dos tecidos com as forças aerodinâmicas. Segundo Kreiman e Sidtis (2011, p. 44), quando as pregas oscilam, interrompem periodicamente o fluxo de ar proveniente

dos pulmões, criando mudanças de pressão no ar por sua abertura e fechamento. Uma vez que os músculos adutores fecham as pregas vocais, há o aumento da pressão do fluxo de ar proveniente dos pulmões que vencem a rigidez das mesmas e forçam sua abertura. Quando a pressão é fraca abaixo das pregas, ou estas estão muito rígidas, elas não abrem; porém, se não estiverem rígidas, abrem e permanecem abertas. Depois que as pregas estiverem abertas, duas forças entram em ação para fechá-las: a primeira é a elasticidade, que as faz voltar a sua configuração inicial assim que a pressão diminui; e a segunda contribuição aerodinâmica para o fechamento das pregas são os vórtices que se formam quando a corrente de ar sai da glote. Os vórtices criam uma pressão negativa adicional entre as pregas, contribuindo para o fechamento mais rápido das mesmas. Uma vez que as pregas estão fechadas, a pressão sob as mesmas aumenta, o que as faz abrirem, e o processo se repete em outro ciclo da fonação. A vibração das pregas vocais modula a corrente de ar proveniente dos pulmões, criando nela um padrão de compressão e rarefação, que se alternam, e ao atravessar a glote, essa vibração se transforma em som. Assim, o período de um som corresponde à sua frequência, de maneira que quanto mais rapidamente as pregas se fecham, maior quantidade de energia de alta frequência é gerada, devido à mudança abrupta de pressão.

A maneira na qual as pregas vocais vibram é responsável em parte pela complexidade acústica dos sinais da fala. A seguir, trataremos dos tipos de fonação resultantes das diferentes dinâmicas verificadas na vibração das pregas vocais.

3.4 FONAÇÃO

Os possíveis estados da glote têm sido fonte para a pesquisa sobre o papel da fonação em diversas línguas. Tal interesse vem do fato de que nós falantes não somente produzimos vibrações regulares nas mais variadas extensões de *pitch*, como também somos capazes de realizar vibrações irregulares com as pregas vocais que resultam em voz áspera, crepitante ou soprosa. Contudo, ao mesmo tempo em que tais vibrações irregulares podem ser controladas por falantes sem

queixa de distúrbios de voz, elas podem também ser as ações incontrolláveis da glote resultantes de uma qualidade de voz patológica. Assim, o estudo interlinguístico dos tipos de fonação revela, segundo Ladefoged (1983), que a patologia de voz de um falante pode ser o fonema do outro.

3.4.1 Tipos de fonação

Variações no modo em que as pregas vocais vibram podem produzir mudanças perceptíveis na maneira como nossa voz soa. O tipo de fonação mais frequentemente produzido por nós seres humanos é a voz modal, resultado da ação relativamente periódica das pregas vocais. Ladefoged (1971) propôs um modelo simplificado de representação dos tipos de fonação por um contínuo de constrição glotal. Nesse modelo, o tamanho da glote, que depende da distância entre as pregas, pode variar do ponto em que elas estão em abertura máxima (como em uma fricativa glotal) até o ponto em que estão completamente fechadas (como em uma oclusiva glotal). Em ambos os extremos, não há fonação. Nos estágios intermediários de abertura entre as pregas, momentos em que há fonação, ocorrem a voz soprosa (*breathy*), voz modal (*modal*) e voz crepitante (*creaky*) (FIGURA 3):

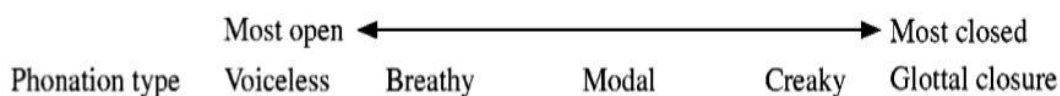


FIGURA 3 - CONTÍNUO DE FONAÇÃO

FONTE: Ladefoged (1971, *apud* Gordon e Ladefoged, 2001, p. 384)

Segundo Keating e Esposito (2007, p. 85), por conta da utilização do contínuo na representação do modelo, é possível conceber a existência de graus não somente de sopro e crepitação, mas também de graus de voz modal. Como consequência, não há categorias discretas na fonação. Além disso, não há correspondência interlinguística na produção da fonação (o que é considerado voz soprosa em uma língua pode ser modal em outra) e nem entre falantes (o que é

considerado voz soprosa para um falante pode ser modal para outro, mesmo dentro da mesma língua).

3.4.1.1 Fonação modal

A fonação modal é caracterizada pela ausência tanto de energia espectral de banda larga nos harmônicos mais altos, quanto de vibração irregular das pregas vocais (Ní CHASAIDE e GOBL, 1997; LADEFOGED e MADDIESON, 1996). Durante a produção da fonação modal, segundo Laver (1991), há tensão adutora moderada e compressão medial das pregas vocais, resultando assim em um vozeamento com periodicidade regular.

Segundo Gerratt e Kreiman (2001), o termo modal teve origem no estudo do registro vocal para o canto, e designa a fonação que compreende a extensão da frequência fundamental normalmente usada para a fala e o canto, ou seja, a distribuição da F0 para um dado falante ou cantor. Fora da literatura sobre o canto, o termo modal é definido conforme a área de estudo: do ponto de vista fisiológico, a fonação modal se opõe à ventricular; do ponto de vista acústico, modal se opõe a outros tipos de fonação que apresentam interharmônicos no espectro; do ponto de vista perceptual, a fonação modal é distinguida da voz crepitante e da diplofonia, por exemplo. Os autores apontam que, por conta do estudo dos tipos de fonação ser conduzido por diversas áreas, a terminologia utilizada nas definições de fonações modal e não-modal acabou se tornando confusa e discrepante. Porém, de maneira geral, a voz modal pode ser definida como uma referência, o tipo de fonação mais usual, em oposição a outros tipos de fonação, tais como a voz soprosa, a voz crepitante e a diplofonia.

3.4.1.2 Fonação não-modal

Quando há alguma variação na vibração das pregas, podemos dizer que se trata de uma voz não-modal. Gordon e Ladefoged (2001) afirmam que controlamos o estado da glote para produzir uma gama de padrões de vibração das pregas – diferentes tipos de *pitch*, aspereza, soprosidade e crepitação. Dentre os tipos não-

modais de fonação, há o falsete, que ocorre nos limites mais altos de frequência do *pitch* da voz de um falante. Nesse tipo de fonação, segundo Kreiman e Sidtis (2011), o contato entre as pregas se dá somente nas suas bordas externas, apresentando um aspecto alongado, fino e rígido, enquanto que as outras partes das pregas permanecem relativamente imóveis. Outro tipo de fonação utilizado na fala é o *period-doubling*, em que há a alternância de ciclos curtos e longos que se repetem, produzindo a sensação auditiva de dois tons simultâneos. Há também a crepitação (*vocal fry*), em que as pregas vocais se abrem e fecham de maneira abrupta, mas permanecem fechadas a maior parte do tempo. Neste trabalho, verificamos a produção dos participantes brasileiros no que concerne aos parâmetros fonatórios modal e não-modais de soprosidade e crepitação, que são a seguir descritos.

3.4.1.2.1 Voz crepitante

A voz crepitante é comumente definida como uma sucessão de excitações laríngeas ou pulsos, de extrema baixa frequência, com abafamento quase completo do trato vocal entre os pulsos (GERRATT e KREIMAN, 2001). Ao contrário da voz modal, a voz crepitante varia de 7 a 78 Hz, independente do sexo do falante, de acordo com Hollien e Michael (1968). O abafamento quase que total do trato vocal entre os pulsos glotais acontece, pois as pregas permanecem fechadas durante grande parte do ciclo e têm curta extensão dos movimentos. O fluxo de ar também é mais fraco do que durante a fonação da voz modal, já que as pregas permanecem fechadas por um período maior de tempo e a extensão do seu movimento é diminuída. Na voz crepitante, há alta tensão glotal, estado em que as pregas vocais estão aduzidas e comprimidas. Perceptualmente, a voz crepitante está diretamente ligada a baixos valores de F0, assim como ao abafamento entre os pulsos, tornando categórica a diferença entre aquela e a voz modal.

3.4.1.2.2 Voz soprosa

A soprosidade é o resultado do excessivo vazamento de ar pela glote quando as pregas vocais não se aproximam totalmente durante a fonação (HANSON *et al.*, 2001). Segundo os autores, para falantes normais, tal vazamento ocorre quando há uma abertura glotal nas cartilagens aritenóide, mantida durante todo o ciclo fonatório. Essa abertura glotal pode resultar em dois efeitos no sinal acústico; primeiro, modifica o componente periódico ao reduzir a amplitude do sinal em frequências médias e altas; segundo, introduz ruído no espectro nessas mesmas faixas de frequências. Para Gordon e Ladefoged (2001, p. 391) a soprosidade pode ser realizada de três maneiras diferentes: (a) pela abertura das pregas vocais; (b) pelo fechamento menos abrupto das pregas; (c) pela abertura posterior das cartilagens aritenoides.

Segundo Gerratt e Kreiman (2001, p. 377), a soprosidade pode ser uma característica pessoal de uma determinada voz, que pode ser avaliada como extremamente soprosa, levemente soprosa, e assim por diante. De acordo com esses autores, na prática, na ausência de um contraste fonológico entre a voz modal e a voz soprosa, fica difícil separar uma da outra, seja sob a perspectiva perceptual, sob a acústica, ou pela configuração glotal subjacente. Portanto, a voz soprosa pode ser tanto a característica de uma voz com patologia, quanto uma qualidade de voz habitual, ou uma característica linguística, a depender da causa subjacente à sua produção.

Hanson *et al.* (2001) apontam que muitos acreditam que as mulheres tendem a produzir a fonação com maior incidência de voz soprosa que os homens, mas relatam que não se sabe se tal diferença é motivada por características fisiológicas ou se é resultado de um construto social.

3.4.1.3 Fonação não-modal alofônica

Segundo Garellek (2012, p. 152), estudos geralmente focalizam a descrição acústica de dados de línguas que utilizam a fonação não-modal de forma

contrastiva²⁶. Contudo, a fonação não-modal alofônica tem um importante papel na marcação de fronteiras prosódicas nas línguas. Gordon e Ladefoged (2001, p. 391) afirmam que a fonação não-modal das vogais é extremamente comum na vizinhança de consoantes que não são produzidas com a fonação modal, como é o caso das vogais que precedem consoantes oclusivas. Segundo esses autores, é também recorrente haver sopro antes de oclusivas não-vozeadas e crepitação antes de ejetivas. A fonação não-modal, especialmente a voz crepitante, é geralmente usada como marcadora de fronteira prosódica, tanto no início quanto no fim da frase entoacional. Ainda no domínio prosódico, há uma grande incidência de voz crepitante no início de frase, principalmente no início de enunciados mais longos e em sílabas acentuadas. Um número relativamente pequeno de pesquisas tem investigado a fonação não-modal alofônica, como é o caso do IN (Ladefoged, 1983; Löfqvist e McGowan, 1992) e do sueco (Gobl e Ní Chasaide, 1999; Löfqvist e McGowan, 1992).

3.4.2 Características acústicas dos tipos de fonação

De acordo com Gordon e Ladefoged (2001, p. 397), há um número de propriedades acústicas, articulatórias e aerodinâmicas que caracterizam as vozes soprosa e crepitante, diferenciando-as entre si e também da voz modal. Para este trabalho, serão focalizadas as características acústicas da voz soprosa e da crepitante, em comparação com a voz modal. Para descrever esses e outros tipos de fonação, Kreiman e Sidtis (2011) afirmam que há dois caminhos: analisar os padrões da voz em função do tempo (medidas de domínio temporal) ou em função da frequência (medidas de domínio de frequência).

No domínio temporal, uma das propriedades acústicas mais importantes da fonação não-modal tem relação com a periodicidade. Como já mencionado na seção 3.4.1.2.1, a voz crepitante apresenta pulsos glotais aperiódicos, ou seja, irregularidades na duração do ciclo glotal que determina a frequência fundamental. A voz soprosa também apresenta aperiodicidade, já que os pulsos glotais são modificados pela presença do ruído na fonação, principalmente em altas

²⁶ Para uma revisão de diversas línguas que usam a fonação de forma contrastiva, ver Keating *et al.*, 2010.

frequências. A Figura 4 mostra a aperiodicidade do sinal acústico pelas figuras dos formatos de onda de vocábulos do Jalapa Mazatec produzidos por uma falante feminina.

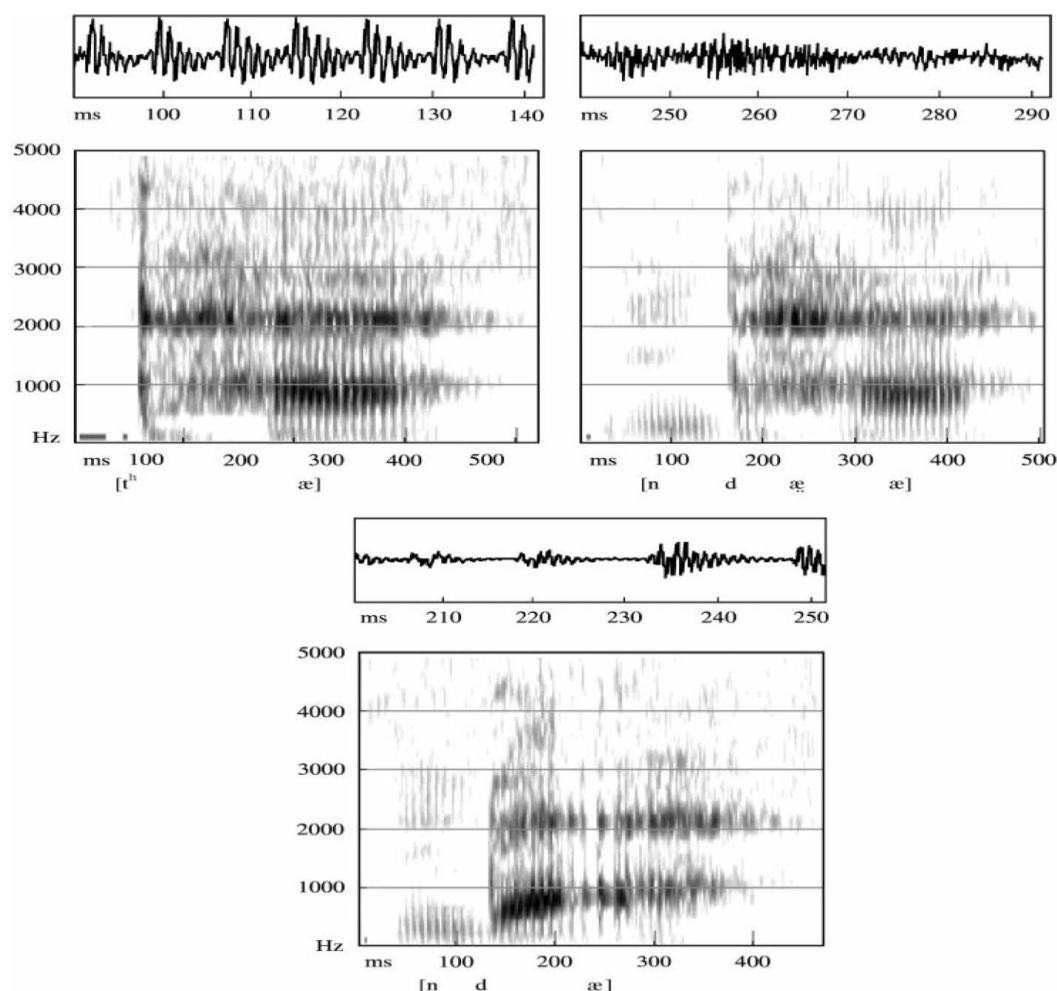


FIGURA 4: FORMA DA ONDA E ESPECTROGRAMA DA FONAÇÃO MODAL, SOPROSA E CREPITANTE, RESPECTIVAMENTE
 FONTE: Gordon e Ladefoged (2001, p. 390).

Com relação à amplitude, ambas crepitação e soprosidade estão associadas à diminuição da intensidade acústica. Como pode ser visto na (FIGURA 4), em comparação à voz modal, esses tipos de fonação apresentam amplitude bastante reduzida. Com relação à frequência, voz crepitante apresenta períodos de *pitch* menos frequentes e irregulares na sua duração.

No domínio da frequência, os tipos de fonação podem ser descritos pelo declínio espectral²⁷, que é o grau em que a intensidade diminui na medida em que

²⁷ Em IN, “*spectral tilt*”.

a frequência aumenta. O declínio espectral é medido pela comparação entre as amplitudes da frequência fundamental (F0, que também pode ser chamada de H1) e dos harmônicos mais próximos do primeiro, segundo ou terceiros formantes. Em comparação com a voz modal, na voz soprosa, há um aumento na amplitude do primeiro harmônico e no declínio espectral. Já na voz crepitante, o declínio espectral é menor, e a amplitude do primeiro harmônico é menor que a do segundo. Isso quer dizer que a queda da energia em frequências mais altas é menor para a voz crepitante do que para a voz soprosa. A Figura 5 traz as diferenças de declínio espectral entre os tipos de fonação, exemplificados por palavras do São Lucas Quivaní Zapotec:

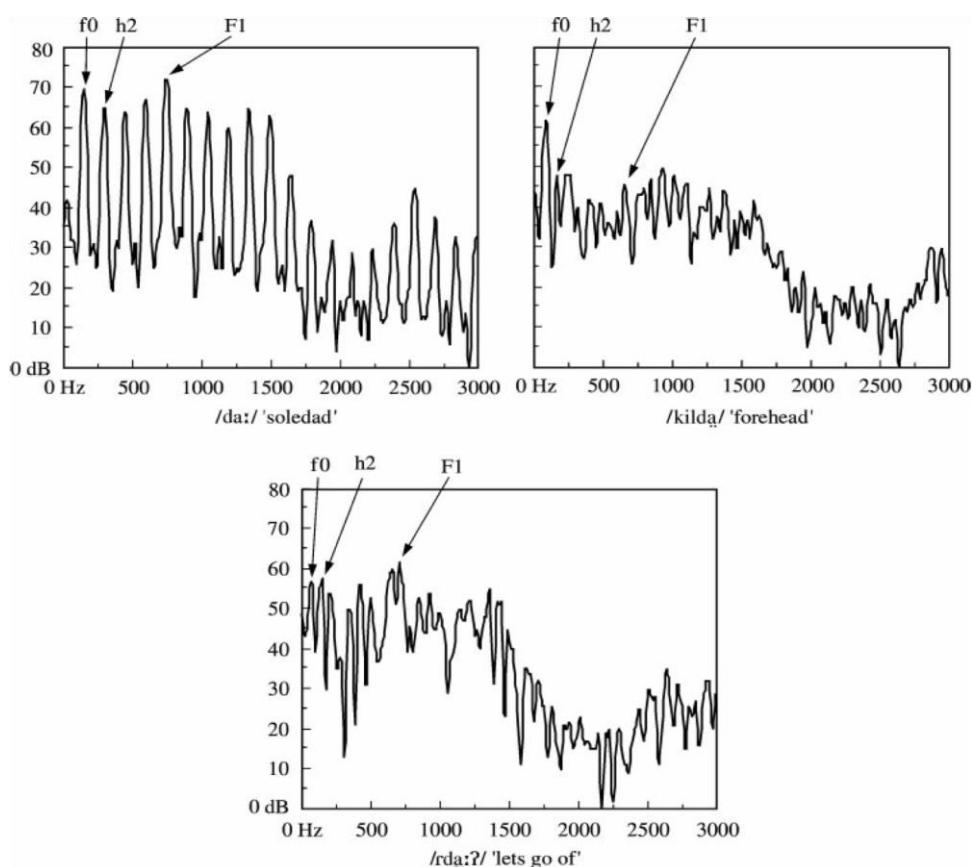


FIGURA 5: ESPECTROS FFT DA VOGAL /a/ NA VOZ MODAL, SOPROSA E CREPITANTE
 FONTE: Ladefoged e Gordon, 2001(p. 398).

Na Figura 5, pode-se ver que no caso da voz modal, as amplitudes da F0 e do harmônico mais próximo do F1 são relativamente parecidas. Já na voz soprosa, a amplitude do harmônico mais próximo do F1 é consideravelmente menor que a

amplitude da F0. Por outro lado, na voz crepitante, a amplitude do harmônico mais próximo do F1 é maior que a amplitude da F0.

A relação entre a F0 e o segundo harmônico também é bastante importante na caracterização da fonação. Como se pode ver na figura 5, na voz modal, o segundo harmônico tem amplitude levemente menor que a da F0. Por sua vez, na voz soprosa, o segundo harmônico tem uma amplitude consideravelmente menor que a da F0. Em oposição, na voz crepitante, a amplitude do H2 é geralmente maior que a da F0.

Tipos de fonação não-modal usualmente apresentam valores menores de F0 em relação à voz modal (GORDON e LADEFOGED, 2001, p. 399). Porém, Keating e Esposito (2007, p. 89) afirmam que tal correlação pode não ser verdadeira para todas as línguas, já que os resultados de algumas pesquisas não encontraram correlação significativa entre valores de F0 e tipos de fonação (HOLMBERG *et al.*, 1989; EPSTEIN, 2002), enquanto outra verificou a correlação entre o aumento do H1*-H2* com o aumento da F0 (ISELI *et al.*, 2006).

A frequência dos formantes também pode ser afetada pelo tipo de fonação. Gordon e Ladefoged (2001, p. 400) apontam que os valores de frequência do primeiro formante da voz crepitante são geralmente mais altos do que os da voz modal e da voz soprosa, fator esse resultante do levantamento da laringe na produção daquele tipo de fonação (SAMELY, 1991 para o Kedang; KIRK *et al.*, 1993, para o Jalapa Mazatec). Por fim, a duração das vogais é outra característica acústica que pode ser alterada pelo tipo de fonação produzido. Kreiman e Sidtis (2011) afirmam que vogais produzidas com ambas vozes soprosa e crepitante têm maior duração do que as vogais produzidas com a voz modal. Gordon e Ladefoged (2001) argumentam que é possível que o aumento na duração das vogais produzidas com fonações não-modais pode acontecer por conta da maior ocorrência desses tipos de fonação em vogais tônicas e não átonas, como por exemplo em Hupa (Gordon, 1998).

3.4.3 Análise dos tipos de fonação

Como afirmam Kreiman e Sidtis (2011), as medidas utilizadas para descrever os tipos de fonação podem ser realizadas sob duas perspectivas: utilizando o domínio temporal ou o domínio da frequência. As medidas feitas a partir do domínio temporal descrevem a proporção entre as fases de abertura e fechamento das pregas, bem como a maneira na qual o fechamento das pregas se dá (se é abrupto ou gradual). As medidas feitas no domínio da frequência correspondem a mudanças de padrão na vibração das pregas e geralmente servem para comparar os níveis de energia em diferentes partes do espectro.

No domínio temporal, um dos modelos mais utilizados para a descrição da atividade glotal é o modelo Liljencrants-Fant (LF), desenvolvido por Fant *et al.* (1985), que apresenta uma proposta de quatro parâmetros para a análise do fluxo glotal. A representação de tais parâmetros se dá por uma derivada da forma de onda do fluxo glotal, cujas partes são frequentemente associadas às fases do ciclo glotal, embora aqueles não tenham necessariamente uma relação direta com a configuração fisiológica da glote (EPSTEIN, 2002).

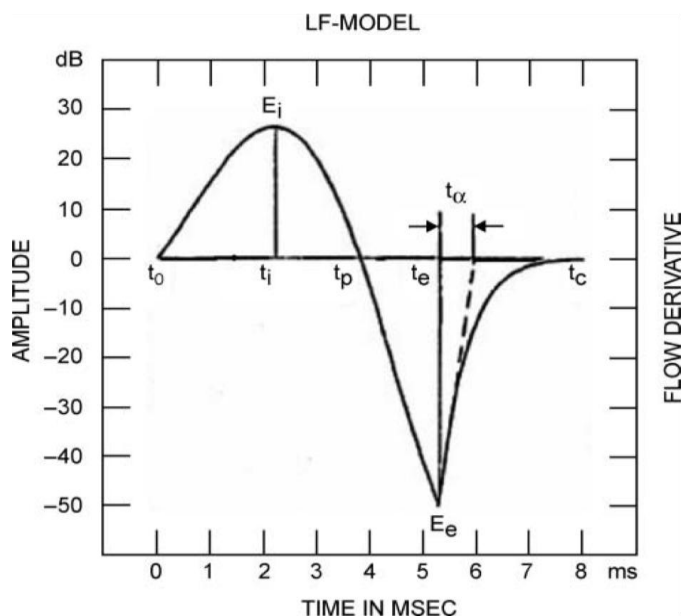


FIGURA 6: MODELO LF DA FONTE GLÓTICA
FONTE: Kreiman *et al.* (2007, p. 2).

Para se obter os quatro parâmetros da análise do fluxo glotal (FIGURA 6), os seguintes aspectos do pulso são utilizados: t_0 , que é o início do pulso; t_c é o comprimento do pulso inteiro; t_p é a duração em que a derivada do pulso glotal é maior que zero; t_i é o tempo em que a derivada glotal atinge seu ponto positivo máximo; E_i é o valor da derivada glotal no seu ponto positivo máximo; t_e é o tempo em que a derivada glotal atinge seu ponto negativo máximo; E_e é o valor da derivada glotal no ponto negativo máximo; t_α é a duração efetiva da fase de retorno. Contudo, para analisar os tipos de fonação por desse modelo, é preciso coletar dados com o auxílio de equipamentos especiais, dificultando, de certa forma, a realização das gravações com tal objetivo. Kreiman *et al.* (2007) apontam que medidas temporais do formato do pulso glotal realizadas pela filtragem inversa somente são precisas se a gravação for feita com a utilização de uma máscara pneumotacográfica ou com um microfone que capture frequências perto de 0 Hz. Além disso, sob a perspectiva perceptual, nem todas as pessoas conseguem ouvir frequências tão baixas, mesmo com o uso de fones de ouvido.

Dada essa relativa dificuldade em realizar a gravação que capture o fluxo glotal, as pesquisas sobre fonação têm recorrido a medidas do espectro, feitas no domínio da frequência, que além de ter um apelo prático bastante grande, são extremamente confiáveis. Iseli *et al.* (2007) confirmam as vantagens a respeito da utilização de medidas acústicas nesse domínio:

Esta abordagem é menos intensa em termos computacionais e menos propensa ao erro que achar a forma da onda do fluxo glotal e é, portanto, ideal para analisar a extensa quantidade de dados necessária para realizar uma avaliação estatística confiável.²⁸ (ISELI *et al.* 2007, p. 2284)

As medidas realizadas no domínio espectral são estimativas da produção da fonação sem a interferência do trato vocal. Tal separação da fonte e do filtro é possível, pois a Teoria Fonte-Filtro de Produção da Fala (Fant, 1960) presume que as duas partes sejam independentes uma da outra. Segundo Alku (2011), para cancelar os efeitos do trato vocal e da radiação labial, um modelo computacional

²⁸ *This is less computationally intensive and less prone to error than finding the glottal flow waveform, and is therefore suited for analyzing the extensive amount of data needed for a reliable statistical evaluation.*

estima, a partir do sinal acústico gravado, o que é a excitação glotal. A (FIGURA 7) ilustra as partes da produção da fala:

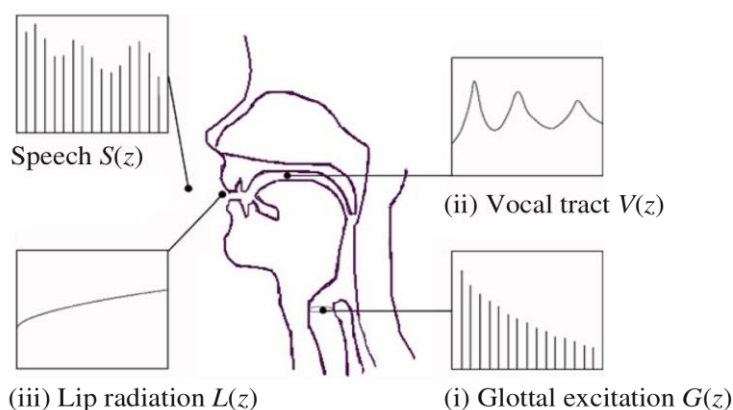


FIGURA 7: MODELO FONTE-FILTRO DE PRODUÇÃO DA FALA
 FONTE: Alku (2011, p. 626).

Segundo a Teoria Fonte-Filtro, representada na Figura 7, a fala é resultado da excitação glotal, que passa pelo trato vocal e é irradiada pelos lábios. O objetivo da filtragem inversa é o de anular os efeitos das fases (ii) e (iii) e chegar à fase (i) pela aplicação de um algoritmo.

Contudo, as medidas espectrais também têm suas limitações e precisam ser feitas com bastante critério. Isso acontece porque tais medidas não capturam os detalhes temporais mais sutis e as mudanças rápidas na fonação – as que acontecem entre um e dois ciclos glotais -, como é o caso da coarticulação entre consoantes e vogais. Por outro lado, por conta da equivalência entre as representações de domínio temporal e de frequência, é possível descrever a relação teórica entre variações temporais na forma do pulso glotal e mudanças correspondentes no espectro da fonte (KREIMAN *et al.*, 2007). Ní Chasaide e Gobl (1997) ressaltam que as medidas no domínio temporal são importantes porque determinam as características espectrais do sinal acústico. Por exemplo, o quociente da abertura (a proporção de tempo em que as pregas estão afastadas em relação ao período do ciclo glotal) e a velocidade de abertura (a proporção de tempo que as pregas levam para se abrir em relação ao tempo que levam para fechar) são variáveis do domínio temporal. Quando as pregas se fecham de maneira não-simultânea no seu comprimento, o bloqueio do fluxo glotal é gradual e

não abrupto, resultando na perda de energia em frequências mais altas e, conseqüentemente, o declínio espectral será maior.

3.4.4 Medidas acústicas do domínio espectral

Kreiman *et al.* (2007) afirmam que, embora se acredite que o declínio espectral seja um parâmetro importante para a caracterização da voz, não há a mesma concordância em se tratando das medidas a serem utilizadas para quantificá-lo²⁹. Algumas pesquisas utilizam o desvio da linha que representa o declínio espectral em relação a um declínio “ideal”. Esse declínio idealizado foi proposto por Carr e Trill (1964) e tem como valor -12 decibéis (dB) por oitava. Porém, como apontam Kreiman *et al.* (2007), até os espectros de vozes normais apresentam variação, já que o declínio não acontece de maneira regular, tornando essa forma idealizada limitada teoricamente. As medidas mais utilizadas são as que representam a diferença entre a amplitude dos harmônicos, que são feitas a partir dos espectros obtidos pela Transformada de Fourier (FFT)³⁰. A análise do declínio espectral pode ser de curto termo, quando as medidas são feitas em vogais sustentadas ou vogais inseridas em palavras de um trecho de fala, ou de longo termo, quando as medidas são feitas em trechos de fala mais longos.

Em análises de curto termo, a medida acústica mais utilizada para distinguir diferentes tipos de fonação é a H1-H2, ou seja, a diferença de amplitude entre o primeiro e o segundo harmônicos. O primeiro harmônico refere-se ao pico de amplitude em dB da f_0 , e o segundo, ao pico de amplitude de duas vezes o valor da f_0 . A relação entre esses dois harmônicos revela características fonatórias da voz. Segundo Hanson *et al.* (2001), é esperado que, à medida em que o quociente de abertura das pregas aumenta, a forma da onda do ciclo glotal se aproxima da

²⁹ Kreiman *et al.* (2007) descrevem outras medidas que não serão utilizadas neste trabalho e, portanto, não serão aprofundadas aqui. Dentre elas estão o Parâmetro Parabólico Espectral (PSP; Alku, Strik, e Vilkman, 1997), que utiliza um único ciclo glotal; outras medidas utilizam a concatenação de um único ciclo ou formam uma sequência de ciclos adjacentes; outra medida é realizada a partir de uma linha de regressão da amplitude dos harmônicos; há ainda o Fator de Riqueza Harmônica (Harmonic Richness Factor; Childers e Lee, 1991), que é a proporção entre a amplitude da F_0 e a soma das amplitudes dos harmônicos de maior frequência.

³⁰ A Transformada de Fourier (*Fast Fourier Transform*) é um método de transformação de um sinal digital que converte o formato de uma onda acústica em um espectro. (JOHNSON, 2012).

forma senoidal, e assim, a amplitude do H1 aumenta em relação aos harmônicos de maior frequência. Dessa forma, H1-H2 tem correlação direta como o quociente de abertura (HOLMBERG *et al.*, 1995). Outros estudos corroboram o fato de que valores altos para H1-H2 estão fortemente relacionados com um quociente de abertura maior (DiCANIO, 2009; KREIMAN *et al.* 2008). Garellek (2012) confirma a importância da medida H1-H2 como um correlato confiável para diferenciar as vozes não-modais da modal, já que a voz soprosa apresenta um maior quociente de abertura enquanto que a voz crepitante envolve a glote mais fechada. Portanto, maiores valores de H1-H2 refletem uma voz mais soprosa, enquanto que se o H2 for mais forte que H1 (produzindo um valor negativo), indica a presença de voz crepitante. Para a voz modal, H1 e H2 apresentam valores similares.

Para que se possa realizar as medidas das amplitudes dos harmônicos no espectro, é preciso também que o efeitos dos formantes e sua largura de banda sejam minimizados. Esse cuidado deve ser respeitado, principalmente se houver comparação entre medidas realizadas em vogais de diferentes qualidades; além disso, em um experimento em que há diversos falantes – que certamente produzem F0 distintas – haverá variação entre a relação da F0 e F1. Hanson (1995) propôs uma fórmula de correção para anular os efeitos de F1 nas amplitudes do H1 e H2 e, mais tarde, utilizou tal fórmula para comparar as vozes de homens e mulheres em vogais não-altas /æ, ε, ʌ/ (HANSON, 1997; HANSON e CHUANG, 1999). Porém, esses dois últimos trabalhos presumem que o valor da F0 está pelo menos 100Hz longe do F1, e não levam em consideração a largura da banda do F1. Posteriormente, Iseli e Alwan (2004) desenvolveram uma fórmula que corrige os efeitos não somente da frequência dos formantes, mas também da sua largura da banda. Esses autores aplicaram a fórmula de correção às vogais /a, i, u/ em estímulos sintetizados. Seus resultados mostraram que a fórmula criada diminuiu consideravelmente a taxa de erro no cálculo das amplitudes dos harmônicos, principalmente no que diz respeito às vogais altas, que apresentam baixo F1. Quando os dados são tratados por essa fórmula de correção dos formantes, a medida H1-H2 é representada com os asteriscos (H1*-H2*).

Além da H1-H2, há medidas que relacionam outros harmônicos do espectro, e que são também relevantes para a caracterização da fonação. A medida H2-H4,

por exemplo, captura alguma característica diferente da qualidade de voz do indivíduo que as outras medidas não o fazem: os resultados de Kreiman *et al.* (2007) mostram que o H2-H4 foi um dos quatro fatores que explicaram a variância em 76,6% dos dados. Baixos valores de H2-H4 são característicos do falsete (quando as pregas estão mais rígidas) e valores mais altos são associados à voz soprosa (quando as pregas estão menos rígidas) (BISHOP e KEATING, 2012, p. 1103). Khan (2012), em um estudo sobre a fonação de vogais do Gujarati – língua que usa a soprosidade de maneira contrastiva – afirma que H2-H4 é um indicador confiável de soprosidade, muito embora ela não estivesse associada a valores mais altos de *pitch*, como é a correlação apresentada em outros estudos.

Há ainda as medidas que também refletem a taxa de declínio espectral em faixas mais altas de frequência. H4-2kHz e 2kHz-5kHz representam o declínio espectral do quarto harmônico ao pico de amplitude mais próximo dos 2kHz, e a diferença entre os picos de amplitude mais próximos de 2kHz e 5kHz, respectivamente. Garellek *et al.* (2013b) afirmam que, especialmente a medida 2kHz-5kHz, está diretamente relacionada ao nível de ruído do sinal acústico, sendo também um correlato da soprosidade. Dessa forma, valores mais baixos dessa medida, indicam presença de voz soprosa, já que quanto maior o declínio, menos ruído no sinal acústico.

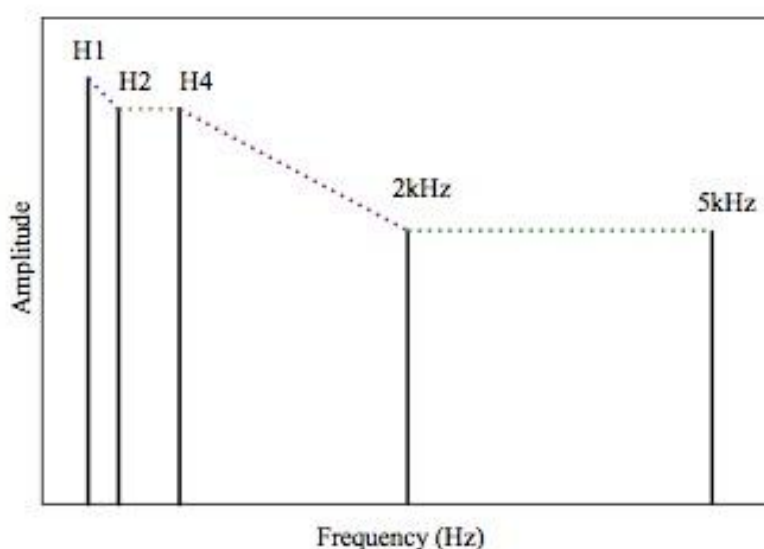


FIGURA 8: ILUSTRAÇÃO ESQUEMÁTICA DO ESPECTRO ESPERADO PARA UM SINAL DE FALA
 FONTE: Garellek *et al.* (2013b, p. 4).

Os autores verificaram que o comportamento do declínio espectral nessas faixas de frequência, ilustrado na Figura 8, se dá da seguinte maneira: se há um declínio de energia acentuado em uma dada faixa, o declínio da faixa imediatamente vizinha terá leve ou nenhum declínio de energia. Como ilustrado na figura 8, há um acentuado declínio entre H1 e H2, mas não há um forte declínio entre os picos H2 e H4; a queda de energia volta a ser acentuada entre H4 e o pico mais próximo de 2kHz, e assim por diante.

Outra medida utilizada para descrever a soproidade na fonação é o pico de proeminência cepstral³¹ (PPC), desenvolvida por Hillenbrand *et al.* (1994). O PPC é uma medida normalizada de amplitude do pico cepstral correspondente à regularidade dos harmônicos. Em um sinal acústico periódico, com estrutura harmônica bem definida, há um componente forte correspondente à regularidade dos picos dos harmônicos. Se não há uma estrutura harmônica bem definida no sinal, então a proeminência desse pico cepstral não será tão forte. O PPC mostrou forte correlação com julgamentos perceptuais de voz soprosa em vogais sustentadas e na fala contínua (HILLENBRAND e HOUDE, 1996).

Quando a análise espectral é feita em trechos de fala mais longos, espectro de longo termo (ELT)³² fornece informações espectrais com relação à distribuição de energia. Mais especificamente, o ELT é resultado da média da amplitude ou da intensidade em uma determinada faixa de frequência de um trecho de fala contínua. De acordo com Li *et al.* (1969, apud Pittam, 1987), o ELT deve ser obtido em trechos de fala entre 30 e 40 segundos, duração suficiente para que os efeitos de segmentos individuais não sejam significativos. Segundo Pittam (1987), pelo fato de a qualidade de voz ser uma característica de longo-termo da fala, as medidas de ELT são bastante apropriadas para descrever tal fenômeno por dois motivos: (a) quando o foco é tratar da qualidade de voz como um marcador social ou de personalidade, e (b) quando julgamentos perceptuais são associados às medidas acústicas. As medidas do ELT têm sido usadas para investigar questões relacionadas à identificação de falantes, comparação entre grupos de culturas

³¹ Em IN, “*Cepstral Peak Prominence*” (CPP). Essa medida é realizada no cepstro, que é o resultado da IFT (Inverse Fourier Transform) do logaritmo do espectro do sinal acústico.

³² Em IN, “*Long Term Average Spectrum*” (LTAS).

diferentes e classificação de estados emocionais. Para Pittam (1987), o ELT é um meio eficaz de obter principalmente informações relacionadas à fonação, como a soprosidade e a tensão laríngea.

Há várias maneiras de medir o ELT: (a) a análise espectral, em que o formato do espectro é descrito com base na média, desvio padrão (DP), assimetria e curtose espectrais; (b) a amplitude do primeiro pico espectral, chamado de nível de pressão sonora³³; (c) a média de energia espectral, medida em relação ao pico mais alto de amplitude; (d) e o declínio espectral, medido pela diferença de energia entre bandas de frequência mais baixa do espectro e as de frequência mais alta. Cada uma dessas medidas está relacionada um determinado aspecto fisiológico da voz ou serve para um desenho de pesquisa específico (LINDENBAUM, 2012). Por exemplo, (a) é uma análise ideal para comparações intra e intergrupos de sujeitos no que concerne às diferenças de características vocais relacionadas à idade, sexo e condições pré e pós-tratamentos. A medida (b) está relacionada com a média de f_0 da voz, e a medida (c), com o valor médio de tensão muscular laríngea que ocorre na fonação (FULLER e HORII, 1988; GOBERMAN *et al.*, 2008). A medida (d) tem relação com a percepção da soprosidade na voz (HILLENBRAND *et al.*, 1994; HILLENBRAND e HOUDE, 1996). Além disso, Hammarberg *et al.* (1980), afirmam que o declínio espectral entre determinadas faixas de frequência do espectro está relacionado à percepção de características fonatórias como a soprosidade e a crepitação.

A medida de declínio espectral de longo termo mais utilizada é a “razão alfa”, que é a proporção de energia entre a faixa de frequência que vai de 0 a 1kHz e a que vai de 1 a 5kHz³⁴. Kitzing (1986) investigou a produção de quatro tipos de qualidade de voz (voz modal em amplitude normal, voz modal em amplitude reduzida, voz sopro e voz áspera³⁵) realizada por fonoaudiólogos experientes. Os resultados dessa pesquisa mostraram que o declínio espectral foi o parâmetro mais potente para distinguir as qualidades vocais produzidas. Hammarberg *et al.*

³³ Em IN, “*sound pressure level*”.

³⁴ Frokjaer-Jensen and Prytz (1976) nomearam essa proporção “*alpha ratio*”.

³⁵ “*Leaky voice*” foi traduzida como voz sopro e “*strained voice*”, como voz áspera.

(1980) também relatam que a diferença de energia entre as faixas mais baixas e mais altas de frequência (como a razão alfa) podem indicar diferentes características glotais, como por exemplo, a ocorrência da voz soprosa quando há uma baixa concentração de energia entre 0,4 e 4kHz. Sundberg e Nordenberg (2006) afirmam que a razão alfa está diretamente relacionada com a intensidade do sinal acústico (quanto maior o volume, mais energia na faixa mais alta de frequência). Com o objetivo de cruzar os resultados de avaliações perceptuais com os de produção da voz, Hammarberg *et al.* (1980) utilizaram a diferença entre os picos de amplitude de diferentes faixas de frequência para relacionar aos resultados da análise fatorial dos julgamentos de percepção. Os autores adotaram as diferenças entre os picos de amplitude de 0 a 2kHz e 2 a 5kHz e entre 2 a 5kHz e 5 a 8kHz. Uma maior diferença entre 0 a 2kHz e 2 a 5kHz (maior declínio espectral) indica a produção da voz soprosa; enquanto que uma maior diferença entre 2 a 5kHz e 5 a 8kHz, é indicativa de voz crepitante.

Portanto, no presente trabalho, utilizamos as quatro medidas propostas por Kreiman *et al.* (2014) – H1-H2, H2-H4, H4-2kHz, 2kHz-5kHz - e o PPC para a análise da fonação produzida pelos falantes bilíngues em vogais do PB e do IN. Adotamos ainda as seguintes medidas de declínio espectral para descrever e comparar aspectos de longo termo: a razão alfa, e as diferenças entre os picos de amplitude em dB de 0 a 2kHz e 2 a 5kHz e entre 2 a 5kHz e 5 a 8kHz. Optamos por essas medidas pelo fato de as pesquisas supracitadas mostrarem evidências que elas estão relacionadas à percepção de diferentes qualidades de voz.

Espera-se que eventuais diferenças de fonação apresentadas pelos participantes do experimento de produção sejam descritas tanto pelas medidas de curto termo (H1-H2, H2-H4, H4-2kHz, 2kHz-5kHz e PPC) quanto pelas de longo termo (razão alfa).

3.5 MEDIDAS DE F0

Sabe-se que línguas e dialetos podem variar no que concerne ao emprego do *pitch*, apresentando diferentes médias e extensões de f0³⁶ (KEATING e KUO, 2012). Segundo essas autoras, se até os falantes de um mesmo grupo social em dada língua podem usar a f0 de maneira distinta, diferenças nesse parâmetro acústico entre línguas ou entre emissões de falantes bilíngues são bastante plausíveis:

A f0 é até certo ponto um aspecto arbitrário da fala, e uma determinada extensão de f0 pode fazer parte da estrutura fonética de uma língua, de maneira que no limite, um falante soaria não-nativo (tendo um sotaque estrangeiro) se usasse uma extensão de f0 diferente³⁷ (KEATING e KUO, 2012, p. 1050)

Se a f0 faz parte da estrutura fonética de uma língua, é possível considerar que um bilíngue utilize diferentes médias ou extensões de *pitch* nas suas duas línguas. Esse pressuposto é baseado em estudos que empregam diferentes metodologias e fatores que potencialmente afetam a produção da f0 em bilíngues, tais como idade, dialeto, experiência na L2 e tipo de tarefa de produção. Alguns estudos sobre as diferenças do emprego da f0 por falantes bilíngues são reportados na seção 4.2.1.

Ao mesmo tempo em que o *pitch* habitual fornece características a uma determinada língua ou dialeto, ele também depende dos aspectos fisiológicos e emocionais do falante como indivíduo. Assim, a f0 depende fortemente de fatores como o sexo, a raça, e características psicossociais do falante. Por essa razão, a análise intra-sujeitos é tão importante quanto a inter-sujeitos na descrição da f0 dos falantes bilíngues, para que os efeitos daquelas variáveis sejam minimizados e a comparação inter-línguas seja confiável.

³⁶ Para maiores detalhes, ver Dolson (1994).

³⁷ *Speaking F0 is to some extent an arbitrary aspect of speech, and a particular F0 range may be part of the phonetic structure of a language, such that in the limit, a speaker would sound non-native (have a foreign accent) using a different F0 range.*

A maioria dos estudos interlínguas para a caracterização da f_0 de falantes bilíngues adota medidas distribucionais de longo termo, ou seja, medidas baseadas na distribuição de valores de f_0 dentro de uma amostra de fala (MENNEN *et al.* 2014). As medidas de f_0 mais utilizadas para comparar emissões de fala bilíngue são a média, o desvio padrão e a extensão, esta última calculada pela diferença entre os valores máximo e mínimo. Tais medidas são influenciadas tanto pelas configurações fisiológicas quanto pelas características habituais de produção e, portanto, podem revelar se o bilíngue realiza diferenças entre sua língua materna e a L2. Portanto, para caracterizar o emprego do *pitch* em PB e em IN como L2 produzido por bilíngues brasileiros, serão utilizadas as medidas de média e extensão da distribuição da f_0 nas amostras de fala obtidas pela leitura de texto e de fala semiespontânea.

A seguir, o modelo psicoacústico para a análise da voz e de sua qualidade adotado no presente trabalho é descrito e discutido, a fim de embasar as escolhas metodológicas quanto às medidas acústicas utilizadas no experimento de produção.

3.6 UM MODELO PSICOACÚSTICO PARA ANÁLISE DA VOZ E SUA QUALIDADE

Questões sobre medidas para avaliar a qualidade de voz têm implicações que vão além do estudo da qualidade em si (KREIMAN e SIDTIS, 2011). Se for possível entendermos a relação entre o sinal acústico e o atributo perceptual utilizado para interpretá-lo, talvez possamos determinar quais parâmetros fisiológicos criam mudanças significativas na fala. Kreiman *et al.*, (2014) deram o primeiro passo para o desenvolvimento de um modelo psicoacústico que unifica a análise da produção e da percepção da voz. Nesse modelo, o que se espera responder são as seguintes perguntas: (a) quando a qualidade de voz muda de alguma forma, o que causa tal mudança?; (b) se uma mudança ocorre na produção da voz, qual será o resultado perceptível na qualidade? (KREIMAN *et al.*, 2014, p. 1). O desenvolvimento dessa teoria se fundamenta na seguinte premissa:

Porque a produção, a acústica e a percepção da voz são todas partes do mesmo processo comunicativo, entender a função comunicativa de qualquer um desses aspectos da voz – laríngeo/fisiológico, acústico ou perceptual – requer o conhecimento de como cada estágio interage com os outros na transmissão da informação vocal.³⁸ (KREIMAN *et al.*, 2014, p. 2)

Para alcançar o objetivo de unificar a produção à percepção da voz pelo sinal acústico, o modelo propõe os seguintes passos:

1. Ligar a percepção à acústica pela descrição de qualidade em termos de medidas acústicas perceptuais válidas que se complementam para determinar a qualidade de voz;
2. Ligar a produção à acústica pela definição de quais mudanças fisiológicas na voz produzem mudanças perceptíveis no sinal acústico;
3. Repetir o processo até que os dois conjuntos de parâmetros acústicos se alinhem.

As mudanças acústicas são avaliadas por sua saliência perceptiva; as que os ouvintes percebem de fato servirão de base para gerar as hipóteses sobre quais mudanças físicas têm consequências perceptuais. Os autores afirmam que, ao identificar atributos vocais perceptualmente importantes e depois examinar os formatos do pulso glotal associados a esses atributos, será possível reconhecer os atributos físicos relevantes à comunicação.

O modelo considera que a qualidade de voz seja o resultado de uma interação entre o sinal acústico e o ouvinte, e que os ouvintes percebem a qualidade de voz como um padrão na sua totalidade, e não como a soma de vários parâmetros. Assim, nem um parâmetro acústico ou um atributo perceptual atrelado a determinada qualidade de voz podem ser avaliados sem o contexto global fornecido pela voz como um todo. A seleção dos parâmetros é feita por dois critérios – incluir parâmetros suficientes para modelar a qualidade de voz, e incluir somente os parâmetros aos quais os ouvintes sejam sensíveis. Para responder às

³⁸ *Because voice production, acoustics, and perception are all parts of the same communicative process, understanding the communicative function of any of these aspects of voice—laryngeal/physiologic, acoustic, or perceptual—requires knowledge of how each stage interacts with the others in the transmission of vocal information.*

perguntas que o modelo se propõe a tentar resolver, os autores vêm realizando uma série de pesquisas, descritas no Quadro 2:

Objetivo	Metodologia	Resultados	Referência
Determinar os parâmetros que variam entre falantes	Análise de componentes principais - diferenças de amplitude entre faixas de frequência no espectro	Os fatores que deram conta da variância foram: o declínio espectral acima de 4kHz, abaixo de 450Hz e entre 1,5kHz e 4kHz; H1-H2 e H2-H4 ³⁹ , declínio espectral e ruído de alta frequência.	Kreiman, Gerratt, e Antofianzas-Barroso, 2007a
Avaliar se os ouvintes percebem diferenças manipuladas nas diferentes faixas de frequência do espectro; medir a eficácia do modelo	Cópia sintetizada de centenas de vozes; teste de percepção igual/diferente no formato AX	Remoção do parâmetro H4-5kHz e inclusão de dois outros: declínio espectral entre H4 e o pico mais próximo de 2kHz, e o declínio entre este último pico de amplitude e o mais próximo de 5kHz.	Kreiman <i>et al.</i> (2010)
Determinar o quão sensíveis os ouvintes são às mudanças aplicadas aos parâmetros	Definição da Just-noticeable difference – JND ⁴⁰ , cópias sintetizadas de 144 vozes, Protocolo “one up, two down” ⁴¹ em um teste de percepção de discriminação de formato AX.	Ouvintes conseguem perceber mudanças menores do que a variabilidade detectada na produção dos falantes e, portanto, o conjunto de parâmetros passa no teste no quesito “necessário”.	Kreiman, J., Garellek, M., Samlan, R. A., and Gerratt, B. R. (2014b).

QUADRO 2: PESQUISAS PARA A CONSTRUÇÃO DO MODELO PSICOACÚSTICO DE PRODUÇÃO E PERCEPÇÃO DA VOZ.

FONTE: Baseado em Kreiman *et al.* (2014).

O primeiro estudo do Quadro 2 é a utilização de uma ferramenta estatística para definir as faixas de frequência do espectro que apresentam maior variância nos dados. Os outros dois estudos referem-se à aplicação de testes de percepção para verificar se os ouvintes são sensíveis às mudanças realizadas nessas faixas

³⁹ Essas medidas estão explicadas na seção 3.4.4.

⁴⁰ Proporção da menor diferença possível em um parâmetro que os ouvintes conseguem detectar de forma consistente.

⁴¹ Quando os ouvintes percebem a diferença entre duas vozes, provocada pela manipulação dos parâmetros, o próximo *trial* contém uma diferença ainda menor entre os estímulos; se os ouvintes julgam erroneamente que os estímulos são iguais, a diferença entre eles aumenta; os trials se repetem até o ponto em que a diferença oscile próxima do valor da JND.

de frequência. Para dar conta da variância da produção da voz, os autores do modelo apresentam evidências obtidas por outros estudos (QUADRO 3):

Evidência	Resultados das pesquisas	Referências
Aspectos da produção da voz que os falantes conseguem manipular para produzir mudanças perceptíveis na qualidade de voz - (uso linguístico da qualidade de voz).	Diferenças entre H1-H2 (ou H1*-H2*) e H2-H4 aumentaram a probabilidade da percepção da soproidade em línguas que utilizam a fonação de maneira contrastiva (White Hmong, Gujarati, Mazatec, Chong e Green Mong).	(Garellek <i>et al.</i> , 2013a; Andruski e Ratliff, 2000; Blankenship, 2002; Fischer-Jørgensen, 1967; DiCanio, 2009; Garellek e Keating, 2011)
Relação entre o quociente de abertura das pregas, obtido por eletroglotografia, e por medidas acústicas.	Distinção entre voz modal, sopro e crepitante pela relação entre o quociente de fechamento e o valor de H1*-H2*.	Esposito (2012)
Mudanças na configuração glotal com mudanças de qualidade de voz em um contínuo entre sopro a constricta.	O quociente de abertura, o coeficiente de assimetria e a f0 deram conta de uma média de 74% da variância de H1*-H2*.	(Kreiman <i>et al.</i> , 2012)
Conseqüências perceptuais e acústicas produzidas por mudanças fisiológicas. O estudo utilizou um modelo mecânico auto-oscilante das pregas vocais para verificar as incompatibilidades de rigidez entre a prega esquerda e a direita.	Mudanças no grau de incompatibilidade de rigidez entre a prega esquerda e a direita e a assimetria de vibração resultante não produziram diferenças significativas na qualidade a menos que a incompatibilidade fosse grande a ponto de mudar modo de fonação.	Zhang <i>et al.</i> (2013)
Relação entre medidas cinemáticas, acústicas e perceptuais usando amostras de voz geradas por um modelo computacional de pregas vocais acoplado a um modelo de trato vocal.	Clara relação entre a medida CPP (Cepstral Peak Prominence), separação dos processos vocais e os julgamentos de soproidade, bem como o abaulamento das pregas e o declínio espectral.	(Samlan e Story, 2011;; Samlan, Story, e Bunton, 2013)

QUADRO 3: EVIDÊNCIAS PARA A CONSTRUÇÃO DO MODELO PSICOACÚSTICO DE PRODUÇÃO E PERCEPÇÃO DA VOZ.

FONTE: Baseado em Kreiman *et al.* (2014).

A partir desses resultados apresentados no Quadro 3, ficou ainda mais evidente que a fonação é parte imprescindível do processo de comunicação, cujos pedaços são difíceis de compreender de maneira isolada. Dessa forma, entender como os falantes produzem a qualidade de voz intencionada requer um modelo que relacione a produção à percepção. Para chegar a essa relação, é necessário mapear as representações temporais e espectrais da fala, o que tem provado não

ser fácil, já que a fonação acontece no domínio temporal, mas a percepção depende fortemente dos aspectos espectrais do sinal acústico. Outra dificuldade está no fato de que mais de uma configuração física produz uma mesma qualidade de voz, ou de mudanças na configuração física que não produzem alteração alguma na qualidade. Por fim, talvez a maior dificuldade esteja na extrema complexidade do funcionamento do sistema fonador, agravada pela dificuldade de acesso para fins de observação e mensuração da fonação, o que acarreta na impossibilidade de um modelamento fiel ao processo como um todo. O modelo de Kreiman *et al.* (2014) ainda não está pronto, já que essas questões ainda precisam ser resolvidas, mas dá um passo importante em direção a uma abordagem sistemática que descreve como as características da produção da voz interagem com os atributos perceptuais para a transmissão de informação na comunicação.

Alternativamente, para capturar a interação entre produção e percepção da qualidade de voz, tem sido comum a utilização da combinação entre o julgamento perceptivo-auditivo e os métodos instrumentais por meio do uso de ferramentas estatísticas para entender a correlação entre os dois tipos de análise. Maryn *et al.* (2009) analisam o uso das medidas acústicas e sua relação com avaliações perceptuais e apontam que, por meio de gravações digitais e análises automatizadas utilizadas hoje em dia, tornou-se possível uma avaliação mais precisa da qualidade de voz. Esses autores relatam que, para relacionar os resultados da avaliação perceptivo-auditiva aos obtidos por medidas acústicas, pode ser usado o coeficiente de correlação, que mede a relação linear entre variáveis dependentes (neste caso, a avaliação perceptivo-auditiva) e independentes (as medidas acústicas).

Na tentativa responder às perguntas da presente pesquisa, adotamos as medidas acústicas utilizadas pelo modelo psicoacústico de produção e percepção da voz (KREIMAN *et al.*, 2014), bem como medidas de espectro de longo termo e de f_0 , para a descrição da voz dos bilíngues no PB e no IN. Aliamos a essa análise instrumental, a avaliação perceptivo-auditiva dos dados produzidos pelos brasileiros em PB e em IN, e utilizamos testes estatísticos para capturar possíveis correlações entre a produção da voz e a percepção de sua qualidade.

3.7 CONCLUSÃO DO CAPÍTULO

Neste capítulo foram apresentadas as definições relacionadas à produção da voz e à percepção de sua qualidade. Nesse sentido, optou-se por adotar os termos “voz” como sinônimo de fala e de “qualidade de voz” como a resposta perceptual de um ouvinte a um estímulo de fala. Para a integração da análise da produção da voz e da percepção da sua qualidade, foi apresentado um novo modelo teórico, proposto por Kreiman *et al.* (2014), que se propõe a testar a produção e a percepção da voz como um fenômeno psicoacústico que integra esse dois domínios, considerados indissociáveis pelos autores. Além disso, foram descritas as características fisiológicas da produção da voz e os principais tipos de fonação utilizados nas línguas. Ademais, foram relatadas as medidas adotadas na literatura para a análise dos tipos de fonação e, dentre essas, as escolhidas para compor a análise do presente trabalho.

Para investigar a produção da fala e a qualidade de voz de falantes bilíngues de PB e IN, serão discutidas algumas questões relacionadas à aquisição de L2, tema do próximo capítulo. Serão levadas em conta as principais variáveis que influenciam o processo de aquisição da fala em L2. Além disso, serão revistos alguns estudos que investigaram a produção da voz em falantes bilíngues, utilizando medidas acústicas discutidas no presente capítulo.

4 AQUISIÇÃO DAS CARACTERÍSTICAS PROSÓDICAS DA L2

Sabe-se que para dominar a pronúncia de uma L2 é necessário adquirir não somente as características segmentais dessa língua, mas também as características prosódicas da mesma. Mennen e Leeuw (2014) afirmam que:

a pronúncia efetiva em L2 envolve não somente aprender a produzir de forma autêntica todos os sons individuais da língua-alvo mas também a aquisição da prosódia da L2, como sua entoação, acento, ritmo, tom e tempo⁴². (MENNEN e LEEUW, 2014, p. 183)

Contudo, como aponta Lengeris (2012) o ensino da prosódia da L2 tem sido negligenciado em sala de aula por uma série de razões, mas principalmente por consequência da popularidade da abordagem comunicativa nas últimas décadas⁴³. Há também a crença de que características prosódicas da L2 não são ‘ensináveis’ e que dificilmente podem ser internalizadas pelos aprendizes depois de uma certa idade. Outro fator que dificulta o ensino da prosódia da L2 é o fato de que muitos professores não são nativos e, por isso, lhes falta a confiança ou a capacidade de reproduzir os padrões prosódicos da língua-alvo.

A dificuldade em ensinar e aprender as características prosódicas da L2 reside principalmente na sua complexidade e nas suas múltiplas funções. Ladd (1996) afirma que as línguas se diferenciam em quatro dimensões:

- (a) dimensão sistêmica, que representa o inventário de elementos estruturais (como o acento de *pitch* e a sinalização de fronteira frasal);
- (b) dimensão funcional, que explica como esses elementos são usados para sinalizar certas funções linguísticas (como foco ou interrogativas);
- (c) dimensão distribucional, que tem relação sobre o quão frequentemente são utilizados os diferentes elementos estruturais e como eles combinam entre si;

⁴² (...)successful L2 pronunciation involves not only learning how to authentically produce all the individual sounds of the target language but also the acquisition of the L2's unique prosody, such as its intonation, stress, rhythm, tone, and tempo.

⁴³ A abordagem comunicativa priorizava a função linguística, deixando a forma em segundo plano. Para mais detalhes, ver Celce-Murcia *et al.* (1996).

(d) dimensão realizacional, ou seja, a implementação fonética dos elementos estruturais.

Assim, um aprendiz de L2 precisa não somente conhecer os elementos estruturais da prosódia da língua, mas também saber como esses elementos são realizados foneticamente, e como eles se combinam entre si para veicular o significado do enunciado. Se por um lado é recorrente a ideia de que aprendizes de L2 têm dificuldades em dominar todas as dimensões apresentadas por Ladd, por outro, não se sabe ainda por que aprendizes normalmente não conseguem produzir uma prosódia adequada da L2. Há muito menos pesquisas que descrevem ou analisam a produção e percepção da fala em L2 no que concerne a aquisição de padrões prosódicos do que padrões segmentais. Baseado em um levantamento feito nos principais periódicos relacionados à aquisição de L2, Gut (2009; apud Mennen e Leeuw, 2014), apontou que, de 133 estudos sobre fonologia, apenas 17 são sobre prosódia. Mesmo assim, nos últimos anos, há um crescente interesse pela pesquisa sobre prosódia no campo da aquisição de L2 e, conseqüentemente, tentativas de desenvolver modelos novos ou já existentes que dêem conta de explicar a aquisição de características prosódicas da L2 (MENNEN e LEEUW, 2014).

Além da complexidade, outra barreira para a aquisição da prosódia pode ter natureza perceptual – assim como na aquisição dos segmentos, a prosódia da L2 pode ser influenciada ou filtrada pelas características prosódicas da L1 do aprendiz. Com base em pesquisas sobre a percepção de tom, acento e entoação⁴⁴, Mennen e Leeuw (2014) relatam que a maioria das dificuldades apresentadas por aprendizes na produção de características prosódicas da L2 parece ser motivada pela falha na percepção de tais elementos. Porém, outros resultados mostram que, apesar de apresentarem uma boa percepção da prosódia da L2, os aprendizes não necessariamente produzem tais aspectos de maneira adequada; da mesma forma, aprendizes podem ter percepção inadequada, porém produção satisfatória da prosódia da L2.

Outra perspectiva adotada por pesquisadores com a intenção de buscar mais informações a respeito da aquisição da prosódia é por meio da avaliação de

⁴⁴ Para as referências, ver Mennen e Leeuw (2014, p. 187-188).

como a produção de padrões prosódicos não-nativos são percebidos por falantes nativos da L2. Nesse sentido, o que se pretende entender é o grau de influência da prosódia na avaliação do sotaque estrangeiro. Há estudos que compararam o efeito de características segmentais com o de características suprasegmentais na percepção do grau de sotaque estrangeiro de aprendizes de L2. Segundo tais pesquisas (Munro, 1995; Munro e Derwing, 1995; Derwing *et al.*, 1998) desvios em características prosódicas tiveram um maior efeito nos julgamentos a respeito do grau de sotaque do que as características segmentais. Embora a maioria desses estudos focalizem o papel da entoação na percepção do sotaque não-nativo, há outras pesquisas que verificaram a influência do alcance de *pitch*, acento, taxa de elocução, duração e ritmo⁴⁵. Os resultados de tais pesquisas são inconclusivos: alguns mostram que os segmentos afetam mais a percepção do sotaque estrangeiro, outros mostram que segmentos e suprasegmentos têm papéis igualmente importantes, e alguns apontam uma prevalência das características prosódicas sobre as segmentais.

Atualmente, sabe-se também que a prosódia está fortemente ligada a questões de inteligibilidade e compreensibilidade da fala (Jilka, 2000; Munro e Derwing, 2001; Trofimovich e Baker, 2006). Um estudo de Tajima *et al.* (1997⁴⁶, apud Mennen e Leeuw, 2014), mostrou que a correção de padrões rítmicos produzidos em L2 por falantes não-nativos aumentou sua inteligibilidade. Munro e Derwing (1995) e Derwing e Munro (1997) verificaram que desvios nos padrões prosódicos produzidos por aprendizes de L2 afetam negativamente sua inteligibilidade e compreensibilidade. Outro estudo mostrou que falantes nativos tiveram menor acesso lexical ao ouvirem contornos de entoação produzidos por não-nativos, reforçando assim, a ideia de que desvios nas características prosódicas afetam a compreensibilidade. Esses achados mostram o importante papel que a prosódia tem na produção e percepção de uma L2, já que ela não somente fornece informação linguística, quanto carrega marcas paralinguísticas da fala.

⁴⁵ Para as referências, ver Mennen e Leeuw (2014, p. 188).

⁴⁶ Tajima, K., Port, R., e Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign accented English. *Journal of Phonetics*, 25, 1–24.

4.1 FATORES QUE INFLUENCIAM A AQUISIÇÃO DE PADRÕES PROSÓDICOS DA L2

A pesquisa sobre produção e percepção em L2 mostra que o processo de aquisição de uma língua estrangeira pode ser influenciado por diversos fatores – sejam eles linguísticos ou extralinguísticos. Segundo Lengeris (2012), idade de aprendizagem, experiência linguística e ambiente de aprendizagem são os fatores mais pesquisados na área de aquisição de L2 de maneira geral. Esses fatores serão discutidos nas seções seguintes.

4.1.1 Idade de aquisição

Quanto mais cedo ocorrer a aprendizagem da L2, melhor – essa é a ideia mais recorrente sobre a influência da variável idade na aquisição da L2, conclusão essa resultante de vários anos de pesquisa. Tal premissa teve como origem o trabalho de Lenneberg (1967⁴⁷; apud Lengeris, 2012), que apresentou o conceito do ‘período crítico’ para a aquisição de L2, referente ao período limite de idade a partir do qual há um declínio na capacidade de aprendizagem devido à mudanças na plasticidade neural. Há evidências que as restrições maturacionais determinadas biologicamente refletem em uma maior dificuldade na produção e percepção de características segmentais da L2, conforme estudos de Flege (1999a) sobre vogais, e de McKay *et al.* (2001) sobre consoantes. No que diz respeito à aquisição de características prosódicas, o estudo de Huang e Jun (2009) examinou o efeito da idade na aquisição da prosódia do IN estadunidense por aprendizes chineses. A variável verificada foi a idade do participante quando da chegada aos Estados Unidos, enquanto que a variável tempo de residência foi controlada. Foram observadas a taxa de elocução, o grau de sotaque e a entoação produzidos em IN pelos chineses. Os resultados mostraram que, em geral, a idade de aquisição tem um efeito na produção dos padrões prosódicos do IN por chineses; porém, a magnitude do efeito variou de acordo com os aspectos da prosódia (taxa de elocução, sotaque estrangeiro e entoação).

⁴⁷ Lenneberg, E.H. 1967. *Biological foundations of language*. New York: Wiley.

Muito embora haja uma série de estudos que forneçam indícios de que a idade de aprendizagem tenha efeito sobre a aquisição fonológica de uma L2, ainda é necessário investigar tal efeito mais profundamente. Lengeris (2012) acredita que mais evidências são necessárias para que se verifiquem três aspectos dessa relação: (a) se existe um declínio súbito na habilidade de aprender uma L2 em uma determinada idade; (b) se todos os aprendizes que adquiriram a L2 quando criança ou muito jovens necessariamente atingem os padrões nativos; (c) se todos os aprendizes que adquiriram a L2 em fase adulta necessariamente fracassam em produzir padrões nativos (LENGERIS, 2012, p. 27). Essas evidências não se confirmam em diversos estudos sobre o assunto: Flege (1995) mostrou que nem todo bilíngue que aprendeu a L2 quando criança apresenta desempenho igualmente bom; Bongaerts *et al.* (1995) verificaram que bilíngues que aprenderam a L2 quando adultos podem atingir uma pronúncia igual a de nativos da língua.

Em conclusão, a questão sobre o efeito da idade de aquisição parece ainda não estar resolvida; há tanto evidências para uma forte relação com a aquisição de padrões segmentais e prosódicos, quanto para a ausência dessa relação. Contudo, a maioria dos estudos considera que, quanto mais cedo ocorra a aquisição, mais bem-sucedida será a produção e a percepção de aspectos segmentais e suprasegmentais da L2.

4.1.2 Experiência linguística

Em pesquisa sobre a aquisição de L2, acredita-se que, à medida que crescemos e nos tornamos adultos, nossa habilidade de aprender uma L2 torna-se mais reduzida por conta da experiência linguística que temos da nossa L1. O Modelo de Aprendizagem da Fala (FLEGE, 1995; 2002) e o Modelo de Assimilação Perceptual (BEST, 1995; BEST e TYLER, 2007) concordam que a experiência da L1 interfere na aquisição da L2, e que a relação entre os inventários de sons da L1 e da L2 prevê quais sons da L2 serão de mais difícil aprendizagem. Estudos sobre características prosódicas, concentrados em avaliar principalmente a aquisição de tom e acento, relatam o efeito da L1 sobre a L2. Mennen (2004) investigou até que ponto a entoação da L1 pode exercer influência na aquisição da entoação da L2 no

nível fonético. Ele examinou a produção do levantamento pré-nuclear do grego por aprendizes holandeses em nível avançado, já que o levantamento ocorre em ambas as línguas, mas têm manifestações fonéticas distintas. Cinco holandeses, todos professores de grego em nível universitário, participaram do estudo. Os resultados mostraram que quatro dos participantes transferiram a característica fonética do holandês na produção do grego.

Embora haja uma quantidade grande de pesquisas que fornecem evidências sobre o efeito da L1 na aquisição da L2, Mennen (2006) afirma que é extremamente difícil comparar resultados de estudos que não compartilham do mesmo encaminhamento teórico e metodológico. Esse autor aponta que as pesquisas geralmente se diferenciam com relação ao nível de proficiência dos aprendizes, número de participantes, diferentes L1 e L2, e metodologia de coleta e análise dos dados. Tais diferenças impedem que os resultados sejam comparados a fim de se estabelecer padrões mais gerais de aquisição das características prosódicas de diferentes L2.

4.1.3 Ambiente de aprendizagem

O ambiente onde ocorre a aprendizagem – em local onde se fala a L2 como língua nativa ou onde a L2 não é falada de maneira naturalística – pode ter um papel importante no desenvolvimento da aquisição tanto de padrões segmentais quanto suprasegmentais da L2. Habitualmente, pesquisas utilizam a variável tempo de residência no país que tem a L2 como língua nativa para verificar seu efeito na aprendizagem. Trofimovich e Baker (2006; 2007) investigaram o efeito do tempo de residência no ritmo, alinhamento de pico, taxa de elocução, frequência e duração de pausas, produzidos por coreanos aprendizes de IN que moravam nos Estados Unidos. As sentenças produzidas pelos participantes coreanos foram julgadas com relação ao grau de sotaque estrangeiro por falantes nativos de IN. No primeiro estudo, os resultados mostraram que todos os coreanos tiveram maior grau de sotaque que o grupo controle (de falantes estadunidenses), e que quanto menor o tempo de residência, maior o grau de sotaque (3 meses *versus* 3 e 10 anos). No segundo estudo, que comparou dois grupos de participantes coreanos

que haviam chegado aos Estados Unidos antes da puberdade, mas variavam o tempo de residência (um ano *versus* 11 anos de experiência) mostrou que os participantes que moravam há 11 anos no país apresentaram pouco ou nada de sotaque estrangeiro em comparação ao grupo que estava no país há apenas um ano.

Mesmo que estudos apresentem evidências de que a imersão tem um impacto positivo na aquisição da L2, Flege (2009) alerta que não somente a quantidade, mas também a qualidade do input são extremamente importantes. Esse autor acredita que a residência em um país da L2 somente traz benefícios ao estrangeiro se este frequentemente engajar-se em situações comunicativas com falantes nativos da L2.

4.1.4 Fatores sociais e afetivos

Para além da idade, experiência linguística e ambiente de aprendizagem, há fatores extralinguísticos que podem afetar a aquisição da L2. São exemplos desses fatores o tipo de tarefa desempenhada pelo aprendiz na coleta dos dados; motivação e atitudes do aprendiz para com a L2; marcas sociais ou de identidade; frequência e qualidade de uso da L2, entre outros. Mais recentemente, estudos têm combinado métodos quantitativos com métodos qualitativos para relacionar o desempenho na L2 com questões etnográficas dos aprendizes. Um destes estudos é o de Moyer (2004), que pesquisou o nível de motivação, a autoavaliação do sotaque, o tempo de instrução formal, e a quantidade e contexto do uso do alemão como L2 produzido por aprendizes de diferentes L1. Os 25 participantes, que eram proficientes em alemão como L2, realizaram quatro tipos de tarefas de produção e foram avaliados por três falantes nativos, além de completarem um questionário que levantava suas experiências pessoais na aquisição da L2. Os resultados da pesquisadora mostraram que, embora a idade tenha influência sobre a aquisição da L2, fatores como a intensidade da motivação, satisfação do alcance do aprendizado em termos de proficiência, e a motivação profissional, foram os fatores que deram conta da maior parte da variância dos dados. Outros fatores que a pesquisadora mostrou serem bastante relevantes para o sucesso na aquisição da

L2 foram: oportunidades de contato; atitude para com a cultura da L2; percepções de pertencimento e estrangeirismo; e intenção de permanecer no país. Dessa forma, a autora aponta para a importância de se pesquisar a aquisição da L2 a partir das diferentes fontes ou domínios de input, e relacioná-los à intenção do aprendiz a respeito da L2.

4.2 QUALIDADE DE VOZ E AQUISIÇÃO DE L2

A qualidade de voz está diretamente relacionada à composição fisiológica do falante e, portanto, tende a ser a mesma independente da língua que a pessoa fale. Por outro lado, a qualidade de voz é também um parâmetro indicial e, segundo Esling (2013), bilíngues ou multilíngues podem alterar a qualidade de voz quando falam línguas diferentes, ou combinar qualidades das línguas que falam. Isso acontece, pois, dentro de uma comunidade linguística, os parâmetros de qualidade de voz são responsáveis por fornecer características tanto regionais quanto sociais dos seus falantes. Além das mudanças na qualidade de voz com motivação dialetal, Esling (2013) afirma haver modulações resultantes da personalidade e do estado emotivo do falante que certamente dependem de padrões culturais de dada comunidade linguística. Da mesma forma, a vontade de modificar dados aspectos da fala em L2 depende da atitude que o falante tem em relação à mesma, além de questões de identidade e autoestima, que são sinalizados pela qualidade vocal.

Jones e Evans (1995) afirmam que diferentes línguas utilizam diferentes modulações de qualidade de voz, o que resulta em uma impressão auditiva de quem fala aquela língua, como por exemplo, a voz labializada e dentalizada do francês ou a voz anasalada ou tensa do chinês. Essa impressão é geralmente o primeiro contato consciente com a fonologia de uma L2, que se é percebido quando alguém imita a maneira como uma língua soa antes mesmo de saber como falar nela. Jones e Evans (1995) defendem que um curso de pronúncia em L2 deve começar pela qualidade de voz, pelo fato de ser uma abordagem holística de diferentes aspectos fonético-fonológicos de maneira integrada. Essa abordagem dá

ao aprender a chance da experiência com a pronúncia da L2 de forma intuitiva, bem como inserida em contextos reais de comunicação.

Segundo Mennen *et al.* (2010), as línguas apresentam características globais resultantes da maneira pela qual usam o trato vocal. Esses autores utilizam o termo “ajuste fonético”⁴⁸ para descrever as tendências articulatórias que o trato vocal mantém em uma configuração habitual de uma língua. Entretanto, descrever as diferenças de ajustes fonéticos entre as línguas tem sido uma tarefa árdua para a pesquisa em aquisição de L2, já que os instrumentos para medir parâmetros articulatórios são poucos e caros. Apesar dessa dificuldade, técnicas experimentais para a investigação articulatória, como o rastreamento de pontos localizados na face (*flesh-point tracking*), a ultrassonografia dos movimentos de língua e a eletropalatografia⁴⁹ podem contribuir sobremaneira para o avanço da pesquisa sobre ajustes fonéticos específicos de cada língua. Um estudo de Wilson (2006), utilizando o rastreamento de pontos na face e ultrassonografia, investigou os ajustes fonéticos de canadenses monolíngues (francês ou IN) e bilíngues. Os participantes monolíngues de IN apresentaram ponta de língua mais elevada, ambos os lábios mais protraídos e com a extensão do estiramento diminuída que os participantes monolíngues falantes de francês. Já nos bilíngues, que eram considerados falantes nativos nas duas línguas, foi confirmada a maior protrusão labial no IN, mas não houve a confirmação da diminuição da extensão dos lábios ou da elevação da ponta de língua. Ao utilizar a eletropalatografia, Gibbon *et al.* (1993, apud Mennen *et al.*, 2010) analisou o encontro consonantal /kl/ em diversas línguas europeias, e verificou que a coordenação temporal de contato entre os articuladores para essas duas consoantes variava no detalhe fino do movimento entre as línguas. Estudos que fazem uso de técnicas experimentais para a análise articulatória são em pequeno número, já que a maioria dos laboratórios não conta com a aparelhagem que os possibilite, fazendo com que pesquisadores recorram a outras formas de abordar o tema.

⁴⁸ Mennen *et al.* (2010) utilizam o termo “ajuste fonético” seguindo o modelo de Laver (1980; 1984).

⁴⁹ Para mais informações sobre essas três técnicas, ver Mennen *et al.* (2010).

4.2.1 Estudos sobre a produção da voz por bilíngues

A maioria dos estudos sobre a produção da voz em bilíngues utiliza medidas acústicas para comparar os dados de fala. Segundo Mennen *et al.* (2010), as medidas acústicas para a análise da fala têm uma metodologia relativamente bem-estabelecida, são não-invasivas e bastante acessíveis. As pesquisas geralmente comparam a produção nas duas línguas utilizando medidas de f0, intensidade, tipos de fonação e espectros de longo termo.

Bruyninckx *et al.* (1991; 1994) investigaram a qualidade de voz produzida por participantes bilíngues falantes de catalão e espanhol. Foram coletados os dados de 24 participantes (12 homens e 12 mulheres), que foram divididos por dominância em uma das línguas. Os participantes foram solicitados a ler um texto foneticamente balanceado em cada língua, repetindo cada leitura cinco vezes. Medidas de ELT - índice de dissimilaridade SDDD⁵⁰ de 0 a 5kHz – foram utilizadas para verificar possíveis diferenças. Duas análises intrafalantes foram realizadas: interlínguas (amostras de catalão X espanhol do mesmo falante) e intra-língua (amostras de catalão X catalão do mesmo falante). Houve maior variabilidade nos resultados interlínguas do que nos intralíngua. Além disso, houve diferenças estatisticamente significativas entre dois grupos no que concerne às comparações intralíngua – mulheres com dominância no catalão e homens com dominância no espanhol. Os autores relatam que:

Há uma tendência para uma maior variabilidade na qualidade de voz para um dado falante na sua língua dominante, enquanto que o grau de coerência tende a ser maior na língua não-dominante.⁵¹ BRUYNINCKX *et al.* (1991, p. 401)

Harmegnies *et al.* (1991) avaliaram o efeito do tipo de *corpus* na pesquisa sobre diferenças de qualidade de voz em bilíngues catalão/espanhol. Os autores utilizaram quatro tipos de *corpora* para cada uma das línguas (A, B, C e D). Os corpora do tipo A são os mais diferentes do grupo: são constituídos por textos

⁵⁰ *Standard Deviation of the Differences Distribution* (Harmegnies, 1988).

⁵¹ *Thus, there is a tendency towards greater voice quality variability for a given speaker in his or her dominant language, while the degree of voice coherence tends to be higher in the non-dominant language.*

foneticamente balanceados, compostos por uma distribuição equilibrada dos elementos fonéticos de cada língua. Os corpora B, por sua vez, são constituídos de frases contendo elementos fonéticos comuns ao catalão e ao espanhol. Os corpora C contem o mesmo formato dos corpora B, porém com a maior similaridade fonética entre os enunciados nas duas línguas que foi possível conceber. Por fim, os corpora D, já que não era possível criar dois enunciados equivalentes foneticamente em línguas distintas, foram compostas frases tão diferentes das dos corpora C quanto os corpora C são distintos uns dos outros. Participaram do estudo seis bilíngues de catalão e espanhol, sendo três (dois homens e uma mulher) com dominância no catalão, e os outros três (dois homens e uma mulher) com dominância no espanhol. Foram gravadas cinco repetições de cada corpus por cada participante; a coleta foi feita em uma das línguas primeiro, com cinco a dez minutos de intervalo antes da coleta na outra língua, não havendo controle da ordem das línguas na coleta. Foram feitas comparações de ELT de dados intrafalantes, tanto interlínguas (catalão-espanhol) quando intralíngua (catalão-catalão e espanhol-espanhol), com textos dos corpora A, B e C em um primeiro experimento, e com os textos dos corpora C e D em um segundo experimento. Os resultados do primeiro experimento mostraram que, mesmo quando a similaridade fonética dos corpora aumenta, a variabilidade interlínguas do falante permanece a mesma. O segundo experimento, que comparou dois corpora equivalentes em similaridade interlíngua, mostrou que há um efeito maior nos valores do ELT da língua utilizada em relação ao efeito do conteúdo dos corpora. Em suma, os autores argumentam que seus resultados vão contra a ideia de que as mudanças nos ELT em comparação interlínguas sejam motivadas pela variação fonética das línguas.

Todaka (1995) analisou a fala de quatro bilíngues (dois homens e duas mulheres), falantes de IN como L1 e japonês como L2. Os participantes foram submetidos a experimentos para a obtenção de medidas aerodinâmicas, acústicas e eletroglotográficas. O objetivo do estudo era comparar as características laríngeas e respiratórias da produção dos bilíngues, e analisar as características supralaríngeas de acordo com a classificação “tensa” versus “frouxa” de Laver (1980). Os resultados estatisticamente significativos ($p < 0,05$) mostraram que: (a) as

participantes mulheres apresentaram voz mais soprosa em japonês do que em IN; (b) todos os sujeitos apresentaram valores de f0 mais altos em japonês do que em IN; (c) todos os participantes apresentaram uso mais amplo do espaço vocálico em IN do que em japonês. Esses resultados sugerem, segundo o autor, que as diferenças de qualidade de voz entre as duas línguas pesquisadas têm motivação sociocultural.

Altenberg e Ferrand (2006) pesquisaram a média de f0 de mulheres bilíngues - nove falantes de russo e IN e as outras nove, de cantonês e IN. Do grupo bilíngue russo/IN, oito disseram ter o russo como L1, e do grupo cantonês/IN, seis consideraram IN sua L1, enquanto que as outras três reportaram ser nativas nas duas línguas. Houve ainda a participação de dez estadunidenses monolíngues para compor o grupo controle. As participantes produziram fala semiespontânea ao responderem o que haviam feito nas férias de verão que precedeu à coleta de dados. A ordem das línguas foi randomizada de modo que metade das participantes gravou primeiramente em IN e a outra metade em russo ou cantonês. Os resultados mostraram que, para as bilíngues russo/IN, houve diferença estatisticamente significativa, sendo a média de f0 mais alta em russo do que a média em IN. Já para as bilíngues cantonês-IN, não houve diferenças estatisticamente significativas entre as médias de f0 nas duas línguas. Não houve diferenças estatisticamente significativas entre médias de f0 do IN quando os três grupos foram comparados. Com relação aos valores de extensão de f0, quando comparados os dados das duas línguas, houve dois resultados distintos: para o grupo russo/IN, não houve diferenças estatisticamente significativas entre a extensão de f0 do russo e do IN; por outro lado, para o grupo cantonês/IN, houve diferença estatisticamente significativa, sendo a extensão da f0 maior no cantonês do que no IN. Outra diferença foi verificada na ordem das línguas durante a coleta de dados: do grupo russo/IN, aquelas que falaram russo primeiro tiveram maior média de f0 em russo do que em IN se comparado àquelas que produziram a fala em IN antes do russo. As autoras argumentam que o fato de algumas diferenças serem encontradas para o grupo russo/IN, mas não para o grupo cantonês/IN, pode estar relacionado a características linguísticas ou culturais ligadas às línguas. A diferença de extensão de f0 produzida pelo grupo cantonês/IN pode ser

explicada pelo fato de o cantonês ser uma língua tonal. Por fim, a diferença entre a ordem das línguas na coleta mostrou que uma língua pode influenciar outra no que concerne à média de f_0 , como aconteceu com as bilíngues que produziram russo primeiro, afetando sua produção do IN.

Ng *et al.* (2012) examinaram a fala de 40 bilíngues (20 homens e 20 mulheres), falantes nativos de cantonês e proficientes em IN como L2, comparando as características vocais dos participantes nas duas línguas. A tarefa do experimento era a da leitura de dois textos: “*The North Wind and the Sun*” em cantonês e “*The Rainbow Passage*” em IN. Os autores mediram média de f_0 , além de medidas obtidas de espectros de longo termo como o primeiro pico, média de energia e declínio espectrais. Os resultados que foram estatisticamente significativos mostraram as seguintes diferenças: as participantes mulheres tiveram maior média de f_0 em IN do que em cantonês; maior média de energia espectral em cantonês do que em IN para todos os participantes; menor média de declínio espectral em cantonês do que em IN. As medidas espectrais revelaram que os participantes, quando falavam cantonês, faziam-no com maior tensão laríngea e com mais ruído de alta frequência, sendo tal diferença relacionada com a predominância da ressonância posterior característica do cantonês. O declínio espectral também revelou maior produção de soprosidade no cantonês em relação ao IN.

Camargo *et al.* (2013) investigaram a produção de um bilíngue falante de IN como L1, espanhol como L2 e português brasileiro como L3 no que concerne aos ajustes de qualidade vocal e a aspectos de dinâmica vocal. O participante, que tinha 71 anos, foi gravado produzindo fala semiespontânea e lendo textos nas três línguas. A análise perceptivo-auditiva foi realizada por meio do protocolo VPAS-PB (Camargo e Madureira, 2008) e a análise acústica foi feita pelas medidas obtidas pelo script SGExpression Evaluator (Barbosa, 2009). Os resultados mostraram que houve diferença na qualidade e na dinâmica vocais entre as línguas nas seguintes características: supralaríngeas (lábios e língua); laríngeas (fonatória e de tensão); medidas acústicas (f_0 , derivada de f_0 e ELT); tipo de tarefa (leitura ou fala semiespontânea). Além disso, a medida manual de declínio espectral revelou o agrupamento das emissões por língua.

Em suma, os estudos relacionados à produção da qualidade de voz por bilíngues revelam que há certas variáveis que influenciam os resultados de maneira geral. Primeiramente, parece haver um efeito na qualidade de voz em função da dominância linguística, pois há evidências que o bilíngue apresenta maior variabilidade na sua língua dominante. Outro ponto crucial na pesquisa sobre qualidade de voz em bilíngues é o tipo de tarefa que o participante realiza, já que parecer haver diferenças entre dados obtidos por meio de leitura de textos ou amostras de fala espontânea ou semiespontânea. Além disso, a variável sexo indiscutivelmente exerce efeito sobre a produção de diferentes qualidades de voz, principalmente motivadas por diferenças fisiológicas e culturais. Fatores como proficiência, idade de aquisição, ambiente de aprendizagem, tempo de instrução na L2 certamente influenciam a produção dos bilíngues, como já está estabelecido pela literatura em aquisição de L2.

4.3 PERCEPÇÃO DA QUALIDADE DE VOZ EM L2

Reconhecer vozes é uma habilidade desenvolvida muito cedo, a partir do momento em que o bebê identifica a voz da mãe e gradualmente expande seu repertório de vozes familiares. Não há limite de número de vozes que compõe esse repertório, que vai sendo construído ao longo da vida de uma pessoa. Contudo, não somente ouvimos vozes familiares em nosso dia-a-dia; entramos em contato com inúmeras vozes que podem tanto continuar a ser não-familiares, ou podem tornar-se familiares com o tempo. A pesquisa na área da percepção da voz considera que usamos diferentes tipos de processamentos cognitivos para a percepção da voz: (a) o processamento por traços, que pressupõe a decomposição do estímulo em elementos aditivos e relevantes; (b) o processamento por reconhecimento de padrões⁵², que implica a apercepção como um todo (KREIMAN e SIDTIS, 2011, p. 158). Na primeira abordagem, considera-se que o processo

⁵² Baseado na percepção de padrões da teoria da Gestalt. Segundo Engelmann (2002, p. 1), “Basicamente, ao se observar coisas do mundo, observa-se suas formas, ou melhor, suas *Gestalten*. A seguir, pode-se dividir essas *Gestalten* em partes. Porém cada parte será sempre parte daquela Gestalt que lhe deu início e não um elemento constituinte básico.”

envolva a percepção e classificação da presença ou ausência de uma série de características previamente estabelecidas. Na segunda abordagem, certas características do estímulo levam à correta interpretação da interação entre as partes e o todo, que por sua vez é determinado pela natureza intrínseca do estímulo e dependem de contexto, familiaridade e memória. Essa diferença entre as duas abordagens faz todo o sentido para a pesquisa em percepção da voz ou da sua qualidade; a tarefa que o ouvinte desempenha em um experimento de percepção pode exigir que ele faça o reconhecimento do padrão da voz como um todo (tarefas de identificação ou reconhecimento) ou atentar aos detalhes característicos da voz (tarefas de discriminação). As tarefas de identificação e de reconhecimento implicam que a voz seja familiar ao ouvinte, enquanto que a tarefa de discriminação pode ser realizada com vozes não-familiars. Kreiman e Sidtis (2011, p. 159) afirmam que “padrões vocais prototípicos subjazem à percepção de vozes não-familiars, enquanto que padrões de Gestalt únicos são utilizados para o reconhecimento das vozes familiares”⁵³. Neste estudo, utilizamos a tarefa de discriminação, já que os ouvintes não têm familiaridade com as vozes dos falantes, para a comparação das emissões em português e em IN que os participantes produziram no experimento de produção.

Na discriminação de vozes não-familiars, a análise de elementos característicos e a comparação entre eles têm papel fundamental. O ouvinte extrai características acústicas básicas como *pitch*, qualidade de voz e taxa de elocução, além de atributos de outra natureza como ‘inteligência’ ou ‘masculinidade’. Kreiman (1987) afirma que o desempenho de ouvintes normais na tarefa discriminatória sugere que ambas as características gerais e as específicas da voz são utilizadas para comparar vozes não-familiars.

No caso da discriminação de vozes em língua estrangeira, se o ouvinte não conhece o inventário fonético da L2, muito provavelmente terá dificuldades em identificar o que é característica específica do falante do que é característica da língua. Segundo Kreiman e Sidtis (2011, p. 241), há quatro possibilidades para a identificação e discriminação de falantes envolvendo línguas estrangeiras: (1) o falante e o ouvinte têm a mesma língua materna; (2) o falante e o ouvinte falam a

⁵³ ...*prototypical vocal templates underlie perception of unfamiliar voices, while unique Gestalt patterns are utilized for familiar voice recognition.*

mesma língua, porém um como língua materna e o outro como L2; (3) o falante e o ouvinte têm uma L2 em comum; (4) o falante usa uma L2 que o ouvinte não conhece. A princípio, a familiaridade com a língua parece contribuir para um melhor desempenho na discriminação de vozes ou de falantes. No presente estudo, os falantes e os ouvintes são bilíngues, falantes de português brasileiro como língua materna e de IN como L2. Assim, o teste de discriminação das vozes tem as características dos itens (1) e (3).

Estudos sobre a discriminação de vozes de falantes bilíngues geralmente visam a identificação de evidências que separam as características indiciais do falante das características linguísticas da L1 e da L2. Goggin *et al.* (1991) testaram o papel da familiaridade com a língua na identificação da voz. Esses autores realizaram experimentos de identificação com estímulos de dois grupos de bilíngues: IN/alemão e IN/espanhol. Os resultados mostraram que: (a) ouvintes monolíngues, falantes de IN, identificaram as vozes dos bilíngues com mais eficácia quando as amostras eram no IN do que quando eram no alemão; (b) ouvintes monolíngues alemães também tiveram melhor desempenho na identificação do falante nos estímulos em alemão; (c) ouvintes monolíngues, falantes de IN, identificaram melhor as vozes dos bilíngues nas amostras em IN do que nas em espanhol; (d) ouvintes bilíngues IN/espanhol não mostraram padrão de identificação definido, mas a identificação dos falantes que apresentavam sotaque em IN foi menor do que quando a amostra não tinha sotaque. Em conclusão, os autores afirmam que a confiança na identificação das vozes dobra quando o ouvinte compartilha da língua apresentada no estímulo em comparação com quando ouve uma língua estrangeira.

Winters *et al.* (2008) investigaram até que ponto a familiaridade com a língua afeta a percepção de propriedades indiciais da fala. Para tanto, os autores realizaram testes de identificação e discriminação da fala da produção de bilíngues (IN/alemão). Em um experimento, os ouvintes foram treinados a identificar os falantes bilíngues em apenas uma das línguas, e testados na sua habilidade de identificar os mesmos falantes produzindo na outra língua. Em outro experimento, ouvintes discriminavam os falantes bilíngues nas duas línguas por um teste do paradigma AX. Os resultados dessa pesquisa mostraram que há informação

indicial que independe da língua⁵⁴ em quantidade suficiente na fala, de modo que ouvintes conseguem discriminar os falantes bilíngues, não importando qual língua falem. Aparentemente, os ouvintes processam a informação indicial⁵⁵ pelas características *language-dependent* quando eles têm familiaridade com a mesma; por outro lado, realizam tarefas de discriminação baseados nas características *language-independent* quando não conhecem a língua. Assim, os autores concluíram que a identificação do falante pelo ouvinte é feita com base tanto em características dependentes da língua quanto em independentes da língua.

O estudo de Wester (2012) verificou a identificação de falantes bilíngues em três grupos diferentes – IN/alemão, IN/finlandês e IN/mandarim. Os ouvintes, 14 para cada par de línguas, eram estadunidenses monolíngues e realizaram tarefa de discriminação de amostras de fala nas duas línguas, decidindo se as frases eram faladas pela mesma pessoa ou não. Os ouvintes apresentaram altos índices de identificação nos julgamentos IN-IN, bem como em julgamentos L2-L2 (e.g. alemão-finlandês), desempenhando de maneira menos acurada quando os estímulos eram IN-L2. Os autores argumentam que, no caso dos julgamentos L2-L2, os ouvintes não eram ‘distraídos’ pelas características linguísticas do sinal da fala, ao contrário do que acontecia nos pares IN-L2.

A percepção e a discriminação de falantes bilíngues também interessam a pesquisadores na área de desenvolvimento de tecnologia de reconhecimento e síntese de fala. Wester *et al.* (2010), pesquisadores participantes de um projeto internacional para o desenvolvimento de um tradutor de fala que mantivesse as características vocais do falante na versão em L2⁵⁶, relataram a necessidade da realização de pesquisas sobre a avaliação da similaridade da voz da fala bilíngue. As questões que precisam ser investigadas são se o falante soa a mesma pessoa em L1 e L2 e se o ouvinte reconhece a similaridade do falante entre as línguas. Embora os autores admitam que o falante possa soar diferente quando fale línguas distintas, eles acreditam que seja muito provável identificá-lo nas duas línguas em

⁵⁴ “*language-independent*”

⁵⁵ A informação indicial é aquela relacionada à voz do falante como indivíduo.

⁵⁶ O projeto foi chamado de EMIME (*Effective Multilingual Interaction in Mobile Environments*). Para mais informações, ver www.emime.org

comparação com outros falantes. Já a questão da percepção da similaridade pelo ouvinte parece ser mais complicada de responder. Os fatores que influenciam a discriminação e identificação das vozes dos bilíngues, já reportados pela pesquisa na área, são: (a) a familiaridade do ouvinte com a L2; (b) a influência dos fatores linguísticos no julgamento da voz; (c) a dificuldade na separação das pistas *language-dependent* das *language-independent* utilizadas pelos ouvintes. Em suma, os autores afirmam que é de extrema importância que se leve em conta as línguas que o ouvinte conhece e as línguas produzidas pelo falante.

Portanto, o presente trabalho propõe um teste de discriminação da voz de brasileiros bilíngues falando português e IN como L2, realizada por ouvintes compartilham da língua materna e da L2 dos falantes. Nosso objetivo é que os ouvintes atentem para a qualidade de voz dos falantes, na tentativa de detectar possíveis mudanças entre as duas línguas. A familiaridade com as línguas pode facilitar a tarefa, e os ouvintes poderão utilizar tanto pistas *language-dependent* quanto *language-independent* para realizar o teste. Os fatores linguísticos terão influência tanto na L1 como na L2, já que os ouvintes são nativos na primeira e proficientes na segunda.

4.4 CONCLUSÃO DO CAPÍTULO

Em suma, o presente capítulo apresentou as questões mais relevantes sobre a aquisição de padrões prosódicos da L2, principalmente sobre a produção da fala e a percepção da qualidade de voz em L2. Foram revisados os principais fatores que influenciam a produção e a percepção dos sons da L2 de maneira geral – idade, experiência linguística, ambiente de aprendizagem e fatores sociais e afetivos. Por meio da revisão das pesquisas sobre a qualidade de voz em bilíngues, ficou evidente que a maioria dos falantes apresenta diferenças entre a produção da voz na L1 e na L2, a depender de fatores como a dominância de uma língua em relação a outra, o tipo de tarefa utilizado no experimento, e questões culturais, principalmente relacionadas ao sexo do participante. Chegou-se a essas conclusões a partir da utilização de medidas espectrais de curto e de longo termo,

bem como variações de frequência fundamental. No que concerne à percepção da qualidade de voz de falantes bilíngues, o fator que tem um papel fundamental é a familiaridade com a língua. Quando ouvintes conhecem a língua dos estímulos utilizados em testes de percepção, eles se baseiam tanto nas características *language-dependent* quanto nas *language-independent*.

Portanto, o presente estudo pretende descrever possíveis mudanças nas características vocais de brasileiros bilíngues, falantes de PB e IN, com a utilização de medidas acústicas obtidas a partir dos dados do experimento de produção, e os resultados obtidos por um teste de discriminação do experimento de percepção. Os procedimentos adotados nesses dois experimentos são detalhados no próximo capítulo.

5 METODOLOGIA

No presente capítulo, são apresentadas as características metodológicas deste estudo no que concerne aos participantes, aos materiais utilizados e aos métodos de análise dos experimentos de produção e de percepção. Primeiramente, é relatado o estudo piloto conduzido para a testagem da metodologia do experimento de produção. Em seguidas são detalhados os experimentos de produção e percepção.

5.1 ESTUDO PILOTO

Para efeitos de avaliação da metodologia, foi realizado um estudo piloto com base no objetivo principal do trabalho: verificar se há diferenças de produção da voz entre as duas línguas dos participantes. Portanto, foram coletados os dados de quatro participantes bilíngues, todas do sexo feminino, sendo três brasileiras e uma estadunidense. A estadunidense tinha 26 anos na época da coleta de dados e era natural de Louisville – Kentucky. Aprendeu português nos Estados Unidos e falava esse idioma havia cinco anos. As participantes brasileiras eram professoras de IN em uma universidade pública, com média de idade de 47 anos (46 a 48), duas do estado do Paraná (Ponta Grossa e Londrina) e uma de São Paulo (Taubaté). Essas informantes lecionavam IN havia 27,5 anos de média (25 a 30) na época da coleta de dados. Todas aprenderam IN como língua estrangeira depois dos 13 anos de idade e utilizavam-no quase que diariamente.

As quatro informantes foram gravadas em um estúdio com tratamento acústico, e realizaram as tarefas de leitura em uma única sessão. As gravações foram feitas em um computador conectado a uma placa de som M-Audio Fast Track Pro, que por sua vez estava conectada a um microfone multidirecional AKG C300B, em uma taxa de amostragem de 44,1KHz. Foi pedido que as informantes falassem em um volume confortável e mantivessem a distância de 10 cm do microfone.

A tarefa que as participantes realizaram foi a leitura de um texto em PB e sua versão em IN. Primeiramente, as informantes fizeram três repetições da leitura da fábula A Cigarra e a Formiga, e depois, três repetições de *The Ant and the Grasshopper* (ANEXO A). Todas as participantes realizaram a gravação da leitura primeiramente em PB.

Foram realizadas as medidas de ELT (perfil espectral de distribuição de energia de 0 a 5kHz), f0 (média em Hz), H1-H2 (dB) e HNR⁵⁷ (dB). Para obter as medidas de ELT, foi selecionado um trecho de 25 a 30 segundos de duração de cada repetição⁵⁸. Para as outras medidas, foram escolhidas vogais abertas e tônicas nas duas línguas: [a] da palavra cantava, e [ɑ], da palavra *bother*. Todas as medidas foram realizadas no Praat (Boersma e Weenik, 2012).

Os resultados desse experimento piloto, a partir das medidas de ELT, mostraram que algumas participantes produzem algumas diferenças em termos de distribuição de energia em diferentes faixas de frequências nas duas línguas. Duas das participantes, sendo uma delas a estadunidense, produziram perfis espectrais bastante parecidos em PB e em IN. As outras duas brasileiras apresentaram variações de maior distribuição de energia em uma das línguas – uma no PB e outra no inglês – em todas as faixas de frequência. Em termos de f0, todas as participantes produziram maiores médias no IN do que no PB. Já os valores de H1-H2, que têm relação com o quociente de abertura das pregas vocais e, portanto, com a soproidade e a crepitação, foram maiores em PB do que IN, mostrando que as participantes produzem as vogais do PB com mais soproidade que as do IN. Inclusive, uma participante brasileira produziu voz crepitante no IN. Por fim, os valores de HNR mostraram que as três participantes brasileiras produziram voz mais soproosa em sua língua materna que na língua estrangeira.

A partir desse estudo piloto, foram verificadas algumas limitações metodológicas no material e na análise dos dados para verificar se os bilíngues modificam a qualidade de voz quando falam línguas distintas. Primeiramente,

⁵⁷ *Harmonics to Noise Ratio*, que indica a proporção entre a parte harmônica do espectro e o ruído.

⁵⁸ “Em um campo, num dia de verão, uma cigarra saltitava e cantava alegremente com todo seu coração. Uma formiga passou levando consigo uma espiga de milho com grande esforço, e ia em direção ao ninho” em português e “*In a Field, one summer’s Day, a Grasshopper was hopping about, chirping and singing to its heart’s content. An ant passed by, bearing along with great toil, an ear of corn it was taking to the nest*”, em IN.

verificou-se que o corpus escolhido, embora fosse adequado para realizar as medidas de ELT por conter duração de aproximadamente 30 segundos, não continha vogais parecidas e que estivessem em contexto fonético idêntico ou análogo nas duas línguas. Além disso, as palavras do PB e do IN que continham as vogais nas quais foram realizadas as medidas acústicas de fonação não estavam na mesma posição sintática, podendo sofrer alterações distintas por conta da entoação da frase. Outra limitação foi realizar somente a tarefa de leitura de textos, e não colher dados de caráter espontâneo ou semiespontâneo, que podem revelar de forma mais evidente as alterações vocais por conta do uso de diferentes línguas.

Ademais, as medidas acústicas também não foram consideradas suficientes para descrever a produção da voz dos bilíngues. As medidas de ELT podem ser mais exploradas para descrever o declínio espectral nas relações entre diferentes faixas de frequência. Além da média de f_0 , outras medidas são bastante relevantes, como por exemplo a extensão, para revelar o uso do *pitch* nas duas línguas. Para a verificação e descrição dos tipos de fonação, podem ser extraídas as medidas que relacionam outros harmônicos do espectro, e não somente os dois primeiros.

Com relação aos participantes, é preferível que se colete dados de sujeitos do sexo feminino e do masculino, principalmente para que se possa investigar se há diferenças no que diz respeito ao uso de qualidades de voz em função do sexo do participante. Outra questão importante é o de controlar variáveis como idade e profissão, já que a literatura mostra que a voz envelhece também ao longo da vida e pode sofrer alterações fisiológicas se o indivíduo usar a voz como instrumento de trabalho, como é o caso de professores.

Há ainda as variáveis relacionadas ao processo de aquisição da L2 pelo bilíngue. Fatores como idade de aquisição, experiência em país de IN, tempo de instrução e conhecimento de outra língua estrangeira podem influenciar a produção e a percepção do aprendiz ou do bilíngue.

Portanto, a partir desse estudo piloto, foram tomadas decisões metodológicas a fim de minimizar os efeitos das variáveis mencionadas na produção e análise dos dados para o estudo definitivo. As decisões foram

baseadas na literatura sobre a produção e a percepção da qualidade de voz e da fala em L2, aspectos já discutidos nos capítulos 3 e 4 deste trabalho.

5.2 EXPERIMENTO DE PRODUÇÃO

O experimento de produção do presente estudo foi desenhado para investigar a produção da voz na fala de bilíngues brasileiros que falam IN como L2. Primeiramente, foi desenvolvido o *corpus* para ser utilizado na coleta de dados. Além disso, pensou-se em um instrumento de levantamento de dados biográficos dos participantes, bem como sua experiência linguística tanto no PB, mas principalmente na aquisição do IN. Em seguida, foi realizada a coleta dos dados. Após essa etapa, foram realizadas as medidas acústicas nas amostras de fala coletadas.

5.2.1 *Corpus*

O corpus desta pesquisa contou com tarefas de leitura e de respostas a perguntas em PB e em IN. Decidiu-se trabalhar com a leitura e a fala semiespontânea neste trabalho, pois, segundo Kreiman e Sidtis (2011), as habilidades de produção da voz diferem-se substancialmente a depender da tarefa: fala (semi) espontânea, leitura em voz alta, repetição e canto. As tarefas de leitura foram compostas por (a) uma frase-veículo, na qual foram inseridas palavras-alvo, e (b) um texto, contendo 98 palavras na versão em IN e 105 na versão em PB. A fala semiespontânea foi obtida pela resposta pessoal a perguntas previamente estabelecidas pela pesquisadora.

5.2.1.1 Corpus para medidas espectrais de curto termo⁵⁹

⁵⁹ Utilizou-se a expressão “medidas de curto termo” para designar as medidas de fonação realizadas nas vogais das palavras-alvo.

Primeiramente, foi composto o *corpus* para realizar as medidas espectrais de curto termo. Tomou-se a decisão de utilizar uma frase-veículo para inserir palavras-alvo para que se obtivesse a produção de uma vogal específica, dentro de um contexto fonético senão idêntico, muito parecido nas duas línguas. Esse desenho foi utilizado para testar a hipótese de que os bilíngues produzem tipos de fonação distintos nas duas línguas. A literatura recomenda realizar medidas dos tipos de fonação em vogais não-altas. Segundo Hanson e Chuang (1999), vogais baixas são mais utilizadas porque o primeiro formante está distante do primeiro harmônico, e o primeiro e o segundo formantes estão relativamente distantes entre si, facilitando assim a extração das medidas acústicas⁶⁰. Assim, foram escolhidas palavras que fossem iguais ou parecidas no PB e no IN. Sabe-se que não há correspondência biunívoca entre as vogais baixas do PB e do IN, portanto optou-se por selecionar palavras que tivessem /ɑ/ ou /æ/ no IN e /a/ no PB. Para que fosse possível comparar a fonação entre as duas línguas, foram escolhidas palavras que fossem iguais ou muito parecidas quanto à vogal átona e as consoantes adjacentes, dissílabas paroxítonas: $C^1V^1C^2V^2(C^3)$, onde C^1 e C^2 foram controladas, e V^1 era a vogal-alvo. Houve grande dificuldade para encontrar palavras iguais ou muito parecidas nas duas línguas e quando essas não foram encontradas, utilizaram-se também alguns logatomas. Pépiot (2014) também utilizou palavras e logatomas dissílabas em sua pesquisa sobre qualidade de voz de bilíngues franceses que falavam IN como L2, justificando que as palavras teriam que ser o mais similar possível nas duas línguas. Algumas palavras reais do IN eram empréstimos do espanhol; porém, como os dados foram coletados no estado da Califórnia, foi pressuposto que a maioria dos participantes as conheceria. Optou-se também por verificar se os falantes conheciam as palavras-alvo utilizadas na frase veículo. Portanto, foi apresentada a eles uma lista com as palavras em PB e em IN e foi pedido que eles marcassem as palavras cujo significado conheciam. Ao todo, foram selecionadas 32 palavras em cada língua, sendo 19 palavras reais e 13 logatomas em IN, e 20 palavras reais e 12 logatomas em PB. Essas palavras foram

⁶⁰ Mesmo com a utilização do programa VoiceSauce (Shue *et al.*, 2011), que usa um algoritmo para corrigir possíveis efeitos dos formantes sobre os harmônicos, optou-se por controlar a qualidade da vogal em que foram realizadas as medidas dos tipos de fonação. Lotto *et al.* (1997), verificou um efeito do posicionamento do corpo de língua na percepção da fonação – vogais altas são geralmente acompanhadas do levantamento do corpo da língua ou avanço da raiz da língua.

inseridas na frase-veículo “Say ____ to her” em IN, e “Diga ____ pra ela” no PB. Foram elaboradas três listas em cada língua (ANEXO B), contendo as palavras-alvo inseridas na frase-veículo que estavam dispostas em ordem aleatória. Assim, cada participante fez três repetições da mesma palavra-alvo.

5.2.1.2 Corpus para medidas de f0 e de ELT

Para realizar as medidas de f0 e de ELT, optou-se por selecionar um texto que fosse foneticamente balanceado, e que produzisse amostras de fala com duração de pelo menos 30 segundos. A escolha do texto – *The Rainbow Passage*, de Fairbanks (1940) – foi baseada na sua ampla utilização por estudos experimentais na área da fonética em IN (PITTAM, 1987; HILLENBRAND e HOUDE, 1996; CLOPPER e SMILJANIC, 2011; KEATING e KUO, 2012). A partir do texto original em IN, foi feita uma versão⁶¹ para o PB (ANEXO C). Foi selecionado apenas o primeiro parágrafo do texto para a realização das gravações, já que essa primeira parte já supria as condições de duração da amostra. Os participantes gravaram três repetições da leitura do texto em cada uma das línguas.

Além da leitura dos textos, os participantes foram solicitados a responder perguntas, o que gerou as amostras de fala semiespontânea desta pesquisa. Foram apresentadas três perguntas aos participantes, e eles tinham que escolher duas dessas para responder (ANEXO D). Para cada uma das perguntas, havia a sugestão de tópicos para as respostas dos participantes, que poderiam ser utilizadas ou não. As mesmas perguntas que eles respondiam em PB, teriam que responder também em IN.

5.2.2 Entrevista para dados biográficos

Para obter dados biográficos e da aprendizagem da língua estrangeira, foi elaborado um roteiro para a entrevista com os participantes, antes da coleta dos dados (ANEXO E). As perguntas da entrevista visaram levantar dados relevantes

⁶¹ Tradução do texto em inglês feita pela autora do trabalho.

ao aprendizado do IN, ao tempo de experiência em países de IN e ao uso dessa L2. Esses detalhes são variáveis que influenciam a produção e percepção em L2 sobremaneira, como já discutido no capítulo 3. Toda a entrevista foi feita em português e individualmente.

Foram coletados dados como sexo, data de nascimento e cidade/estado natal. Foi perguntado ao participante se ele ou ela cresceu na cidade onde nasceu; em caso negativo, foi perguntado onde o participante cresceu. Como os dados de produção foram coletados nos Estados Unidos, foi perguntado ao participante há quanto tempo ele ou ela estava nesse país. Os participantes foram indagados se tinham experiência em outro país de IN e por quanto tempo; somou-se a esse fator o tempo em que os participantes já estavam nos Estado Unidos. Foram também coletados dados a respeito da idade em que iniciou o aprendizado do IN, onde esse aprendizado ocorreu (escola regular, escola de idiomas, professor particular) e por quanto tempo ele durou. Outro item da entrevista era a respeito da porcentagem de uso do PB e do IN; os participantes relatavam a porcentagem de tempo que usavam sua L1 e sua L2 no seu dia-a-dia. Por fim, os participantes relataram se tinham conhecimento de outra língua estrangeira ou não.

5.2.3 Participantes

Os participantes do experimento de produção foram recrutados na *University of California Los Angeles* (UCLA). Todos eram estudantes de graduação ou pós-graduação nessa universidade. Participaram deste estudo 16 brasileiros bilíngues, a maioria com idade entre 20 e 30 anos na época da coleta dos dados, tendo o PB como língua materna e o IN como L2 (QUADRO 4)⁶².

Participante	Sexo	Idade	Origem
Fal1	M	25	São Paulo – SP
Fal2	M	26	Jaraguá do Sul – SC
Fal3	F	24	Caxias do Sul – RS
Fal4	F	23	Florianópolis – SC
Fal5	F	39	Niterói – RJ

⁶² Os participantes do experimento de produção são chamados de “falantes”. Nas tabelas, serão utilizadas “fal” e o número do participante.

Participante	Sexo	Idade	Origem
Fal6	F	34	São Paulo – SP
Fal7	F	30	Chapecó - SC
Fal8	M	30	Curitiba - PR
Fal9	M	21	Florianópolis - SC
Fal10	F	27	Santo André - SP
Fal11	F	20	São Paulo - SP
Fal12	F	29	Florianópolis - SC
Fal13	M	31	Salvador - BA
Fal14	F	25	Belo Horizonte - MG
Fal15	F	23	Brasília - DF
Fal16	M	33	São Paulo - SP

QUADRO 4: SEXO, IDADE E ORIGEM DOS PARTICIPANTES DO EXPERIMENTO DE PRODUÇÃO.

Os participantes revelaram detalhes sobre a experiência de estudo de IN, bem como a vivência em países em que IN é falado como língua materna e o uso dessa L2 no cotidiano (QUADRO 5).

Participante	Idade início aprendizado (anos)	Tempo de estudo IN (meses)	Tempo em país de LI (meses)	Uso da LI no dia-a-dia (%)
Fal1	10	96	4	70
Fal2	8	96	10	70
Fal3	14	36	24	40
Fal4	11	120	5	75
Fal5	14	36	144	75
Fal6	22	36	60	50
Fal7	18	24	60	60
Fal8	7	120	30	80
Fal9	7	24	5	90
Fal10	12	24	7	95
Fal11	10	108	84	90
Fal12	9	96	48	60
Fal13	8	60	68	90
Fal14	11	168	168	100
Fal15	7	96	84	80
Fal16	12	48	68	95

QUADRO 5: DETALHES SOBRE AQUISIÇÃO E EXPERIÊNCIA DO IN DOS PARTICIPANTES DO EXPERIMENTO DE PRODUÇÃO

O Quadro 5 traz a descrição detalhada dos dados obtidos pela entrevista. Primeiramente, com relação à idade do início da aquisição do IN, a maioria dos

participantes relatou ter começado a aprendizagem quando criança (dos 7 aos 12 anos). Somente alguns relataram ter iniciado a aquisição quando adultos (dos 14 aos 22 anos). Muitos relataram terem estudado inglês formalmente em escola de idiomas no Brasil ou em outros países por vários anos. Quanto ao número de meses que os participantes tinham morado em países de LI, houve uma grande variabilidade no número de meses de permanência. Alguns sujeitos moravam no país há mais de 14 anos, enquanto outros haviam chegado poucos meses antes das gravações acontecerem. Por fim, os participantes relataram a porcentagem do tempo que usavam IN no seu dia-a-dia. A maioria reportou usar o IN em maior proporção que o PB no dia-a-dia.

Ademais, alguns participantes relataram conhecimento em outras línguas estrangeiras, todas aprendidas depois do IN: quatro sujeitos relataram conhecimento em espanhol; um em alemão; um em francês; um em japonês; um em mandarim; e um em coreano. Os outros participantes reportaram não ter conhecimento em outras línguas estrangeiras.

Embora esteja bem estabelecida na literatura sobre L2 a influência dos fatores idade de aquisição e tempo de experiência no exterior, nenhum participante foi excluído, apesar da grande variabilidade nessas variáveis, apresentadas nos Quadros 4 e 5. Essa decisão é justificada pelo número reduzido de sujeitos que se voluntariaram para participar da pesquisa. Não houve ainda como controlar a variável proficiência na IN, já que uma restrição desse tipo acarretaria em um número ainda mais reduzido de participantes possíveis. Assim, para tentar contornar as diferenças de proficiência e do efeito de tempo de estudo e experiência no exterior, foi realizado um julgamento de grau de sotaque estrangeiro por falantes nativos.

5.2.3.1 Julgamento do grau de sotaque dos falantes

Para esse julgamento, foi selecionada uma frase das gravações da leitura do texto em IN de cada participante. Decidiu-se selecionar um trecho da leitura do texto ao invés da fala espontânea para que as amostras fossem o mais homogênea possível, além de controlar o conteúdo semântico e sintático da frase. O trecho

“*There is, according to legend, a boiling pot of gold at one end*” foi escolhido pelos seguintes critérios: (a) por conter segmentos de relativa dificuldade para os brasileiros (por exemplo, [ð, ɪ, ə, ɑ, p^h] e oclusivas em coda) e ter curta duração (5 segundos⁶³). Das três repetições da leitura do texto feitas pelos participantes, foi selecionada a frase que continha a melhor fluência para compor os estímulos.

Participaram desse experimento como ouvintes dez falantes nativos de IN, todos estadunidenses, sendo cinco do sexo masculino e cinco, do feminino. A idade dos ouvintes variou entre 18 a 28 anos (média de 23,44 anos); oito ouvintes eram naturais da Califórnia, das cidades de Long Beach (3), Los Angeles (2), Eureka (1), Whittier (1) e Santa Monica (1), e dos estados de Nova Iorque (1) e Washington-DC (1). Todos eram estudantes da UCLA em nível de graduação ou pós-graduação. Nenhum dos ouvintes falava alguma língua estrangeira com fluência, mas relataram ter conhecimentos básicos de espanhol (5), mandarim (1), japonês (1), francês (1) e italiano (1). Como os participantes foram recrutados no Departamento de Linguística da UCLA, todos reportaram ter conhecimentos sobre fonética e fonologia.

O experimento de percepção foi montado no programa de computador Sky, desenvolvido no *Bureau of Glottal Affairs* (BGA), que faz parte do *Department of Head and Neck Surgery* da *David Geffen School of Medicine* da UCLA. O BGA é composto por um grupo de pesquisadores de diversas áreas⁶⁴ que realiza pesquisa sobre a produção e a percepção da voz humana. O programa Sky foi desenvolvido para realizar experimentos de percepção da voz, e possibilita que os ouvintes realizem o julgamento de amostras de fala uma em comparação a outra, com a utilização de uma escala deslizante, como ilustrado na Figura 9:

⁶³ O programa Sky, que foi utilizado para essa avaliação de sotaque, não permitia a utilização de arquivos muito longos e exigia que os mesmos tivessem exatamente a mesma duração.

⁶⁴ Linguistas, fonoaudiólogos, médicos e engenheiros.

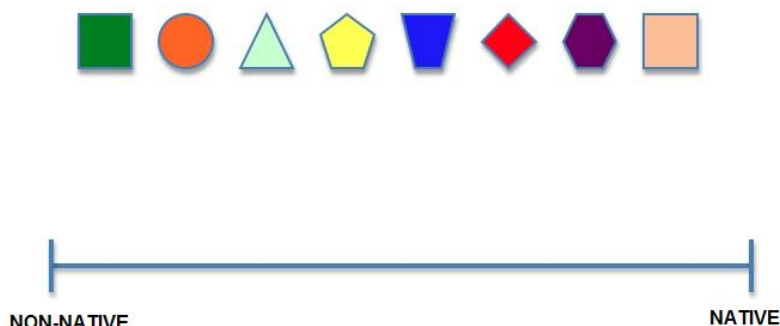


FIGURA 9: REPRESENTAÇÃO DA TELA DA APRESENTAÇÃO DOS ESTÍMULOS NO PROGRAMA SKY

A Figura 9 é uma representação da maneira como os estímulos são apresentados no experimento de percepção para o julgamento de vozes. As figuras geométricas são originalmente coloridas, e são designadas aleatoriamente quanto à forma e à cor, representando cada amostra de fala. A linha, que se encontra na parte de baixo da figura, é um contínuo de mil pontos para a medição do que o pesquisador queira investigar. No caso da presente pesquisa, que tinha por objetivo medir o grau de sotaque estrangeiro dos bilíngues brasileiros, à esquerda do contínuo estava a expressão “*non-native*” e, no outro extremo, “*native*”. Assim, os ouvintes tinham que julgar as amostras de fala dos brasileiros, comparando umas às outras em relação ao quão não-nativo ou nativo o falante soava. As amostras foram apresentadas em dois grupos: um grupo com as vozes masculinas e outro com as vozes femininas. Optou-se por dividir os estímulos em dois grupos por conta de uma limitação do programa, que não permitia que os 16 estímulos fossem apresentados de uma vez só. Para ouvir cada amostra, o ouvinte clicava duas vezes com o botão direito do *mouse* na figura geométrica e esperava até que a duração da amostra se encerrasse para ouvi-la novamente ou ouvir outra amostra. O participante podia ouvir as amostras quantas vezes quisesse. A partir daí, o ouvinte arrastava as figuras geométricas para o contínuo, conforme seu julgamento sobre o sotaque estrangeiro, de forma a dispor no contínuo todas as amostras, umas em relação às outras. Ao final do julgamento, o leiaute da tela do computador ficava como a representação da Figura 10:

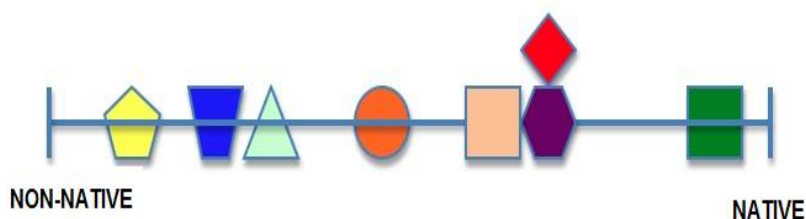


FIGURA 10: REPRESENTAÇÃO DA TELA APÓS O JULGAMENTO DOS ESTÍMULOS PELOS OUVINTES

Os ouvintes podiam arrastar as figuras geométricas e dispô-las no contínuo como preferissem quanto ao espaço entre as amostras, até mesmo considerá-las iguais quanto ao grau de sotaque estrangeiro, como é o caso do hexágono e do diamante. A partir desse diagrama, foram geradas planilhas com os números de zero a mil equivalentes ao posicionamento da amostra no contínuo, de maneira que quanto menor o número, mais sotaque estrangeiro o participante apresentava.

TABELA 1: JULGAMENTO DO SOTAQUE ESTRANGEIRO DOS PARTICIPANTES BRASILEIROS POR FALANTES NATIVOS DE IN

Participante	Média	DP
Fal1	324	211
Fal2	296	197
Fal3	295	254
Fal4	153	129
Fal5	441	226
Fal6	253	290
Fal7	812	139
Fal8	451	253
Fal9	228	179
Fal10	450	166
Fal11	973	15
Fal12	438	256
Fal13	874	104
Fal14	828	69
Fal15	168	149
Fal16	290	146

A Tabela 1 mostra a média e o desvio-padrão do julgamento atribuído a cada participante quanto ao grau de sotaque estrangeiro. O programa utiliza uma escala de mil pontos, com valores mais altos de média indicando participantes com menos sotaque estrangeiro, e valores mais baixos, maior grau de sotaque. Os

valores altos de DP mostram que houve bastante variabilidade nos julgamentos feitos pelos falantes nativos.

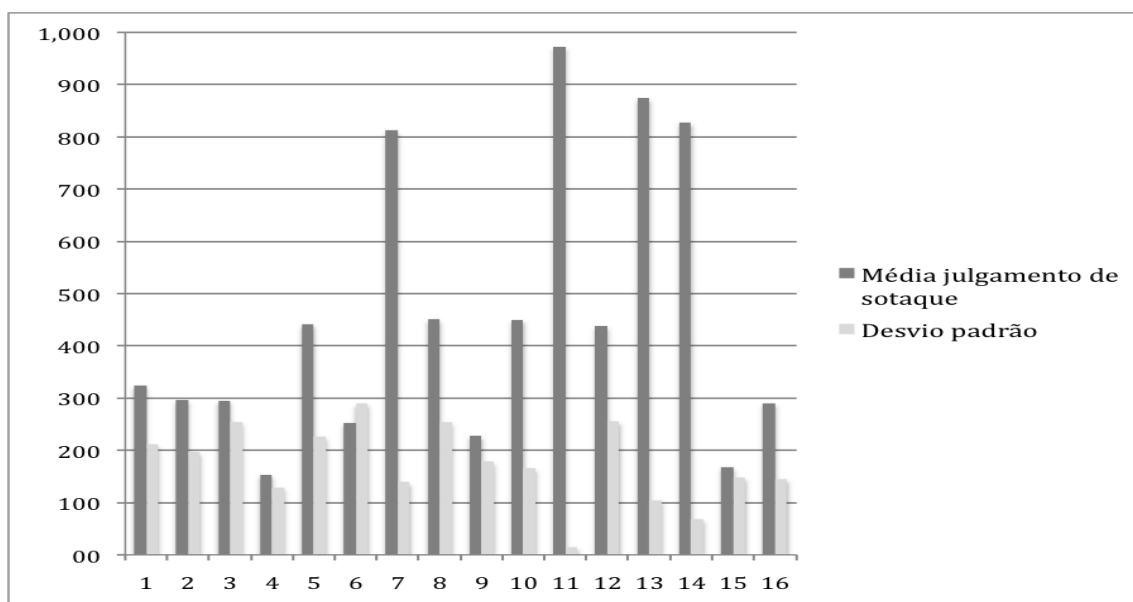


FIGURA 11: MÉDIA E DP DOS JULGAMENTOS DE SOTAQUE ESTRANGEIRO

Ao rodar a análise aglomerativa de k-médias em três grupos, os falantes foram agrupados da seguinte forma: baixo grau de sotaque estrangeiro (Fal7, Fal11, Fal13 e Fal14), médio grau de sotaque estrangeiro (Fal3, Fal5, Fal8, Fal10 e Fal12), e alto grau de sotaque estrangeiro (Fal1, Fal2, Fal4, Fal6, Fal9, Fal15 e Fal16).

Com base nesses julgamentos e nas medidas acústicas realizadas nos dados de produção, pretende-se verificar a relação entre o grau de sotaque e eventuais mudanças na qualidade de voz dos participantes.

5.2.4 Gravações

O experimento de produção foi realizado no BGA, na UCLA, no início do ano de 2014. Primeiramente, os participantes leram e assinaram o termo de consentimento⁶⁵. Depois disso, eles responderam às questões da entrevista, que

⁶⁵ O experimento de produção teve aprovação do Comitê de Ética da UCLA (*Institutional Review Board - IRB*), sob o protocolo IRB#13-001788, a partir da data de 08/01/2014. A coleta de dados foi iniciada dia 13/02/2014. O termo de consentimento está anexado ao final deste trabalho (ANEXO F).

foi feita em PB. Foram gravadas primeiramente as tarefas em PB e, logo em seguida, as do IN. A duração total da coleta foi de aproximadamente 45 minutos para cada participante, sem pausas. As etapas do experimento ocorreram da seguinte maneira:

1. Leitura das frases-veículo em PB. Instruções: ler as frases no volume e velocidade habituais e repetir a frase caso ocorra algum erro. Foi informado aos participantes que havia palavras reais do PB e palavras que não existiam nessa língua. Foi pedido que lessem, em silêncio, a lista de frases antes da primeira gravação. Depois disso, foram gravadas as três repetições da leitura de frases, a partir de três listas diferentes nas quais a ordem das frases foi randomizada.
2. Leitura do texto em PB. Instruções: ler o texto no volume e velocidade habituais. Foi pedido aos participantes que lessem o texto em silêncio para se familiarizarem com o mesmo. Posteriormente, foram feitas as três gravações da leitura desse texto.
3. Resposta às perguntas propostas. Instruções: ler as perguntas e escolher duas para responder. Foi solicitado aos participantes que falassem por cerca de um minuto em cada resposta.

Após a coleta em PB, foram realizadas as mesmas etapas na coleta dos dados em IN: leitura das frases-veículo, leitura do texto e resposta às perguntas. Foi solicitado que o participante respondesse em IN as mesmas perguntas que havia respondido em PB. Depois das gravações, foi apresentada uma lista com as palavras-alvo em PB e em IN contidas na frase-veículo, e solicitado ao participante que marcasse as palavras cujo significado conhecia. Todas as instruções dessa etapa foram as mesmas utilizadas na coleta dos dados em PB, porém dadas em IN.

A gravação dos dados de fala foi coletada em um computador da marca *Apple*, modelo *MacBookPro*, com processador *Intel Core i5* de 2,5GHz, com memória de 4GB e capacidade de armazenamento de 500GB. Foi utilizado um microfone de cabeça, unidirecional, marca *Shure*, modelo SM10A-CN, com resposta de frequência entre 50Hz e 15kHz. Esse microfone foi conectado a um pré-amplificador, marca *Blue*, modelo *Icicle*, que por sua vez foi conectado ao

computador. O microfone foi posicionado a 10 centímetros da boca do participante, em um ângulo aproximando de 45°. O software utilizado para gravação foi o *Audacity 2.0.5*⁶⁶, que é um programa gratuito para edição de áudio. O registro dos dados foi feito em formato *wav*, com frequência de amostragem de 44kHz e taxa de quantização de 32 bits.

5.2.5 Realização das medidas acústicas

Como descrito no capítulo 3, foram escolhidas algumas medidas para investigar o tipo de fonação produzidos em PB e em IN, e outras medidas para examinar a *f0* e o ELT das amostras de fala dos participantes. As medidas acústicas de declínio espectral para os tipos de fonação foram: H1*-H2*, H2*-H4*, H4*-2kHz, 2kHz-5kHz e o PPC. Essas medidas foram efetuadas nas vogais-alvo das palavras inseridas na frase-veículo. Para as tarefas de leitura dos textos e fala semi-espontânea, foram geradas medidas de *f0* (média, extensão e desvio padrão), e o declínio espectral pela razão alfa e pela diferença entre os picos de amplitude das faixas de frequência 0 a 2kHz e 2 a 5 kHz, e de 2 a 5kHz a 5 a 8kHz.

5.2.5.1 Medidas espectrais de curto termo

As medidas de fonação foram realizadas na vogal /a/ dos dados do PB e nas vogais /a/ e /æ/ das palavras em IN. Com a utilização do *software* livre *Praat*, versão 5.3.82 (Boersma e Weenik, 2014), foi segmentada a porção central das vogais-alvo seguindo dois critérios de seleção: (a) pelo menos três pulsos completos e regulares e (b) duração entre 20 e 35 milissegundos. Assim, foram evitadas as bordas da vogal, onde se verifica a coarticulação dos segmentos adjacentes, então selecionou-se a parte central pela inspeção visual da estabilidade formântica. De acordo com Khan (2012, p. 10), que investigou a eficácia de medidas acústicas para descrever as vogais do gujarati⁶⁷, a voz modal e

⁶⁶ Disponível em <http://audacity.sourceforge.net>

⁶⁷ O gujarati é uma língua indoariana que usa a fonação soprosa de maneira contrastiva.

a sopro não são distintas nas suas bordas, mas sim, na sua porção central. Um exemplo da seleção do trecho da vogal pode ser visto na Figura 12):

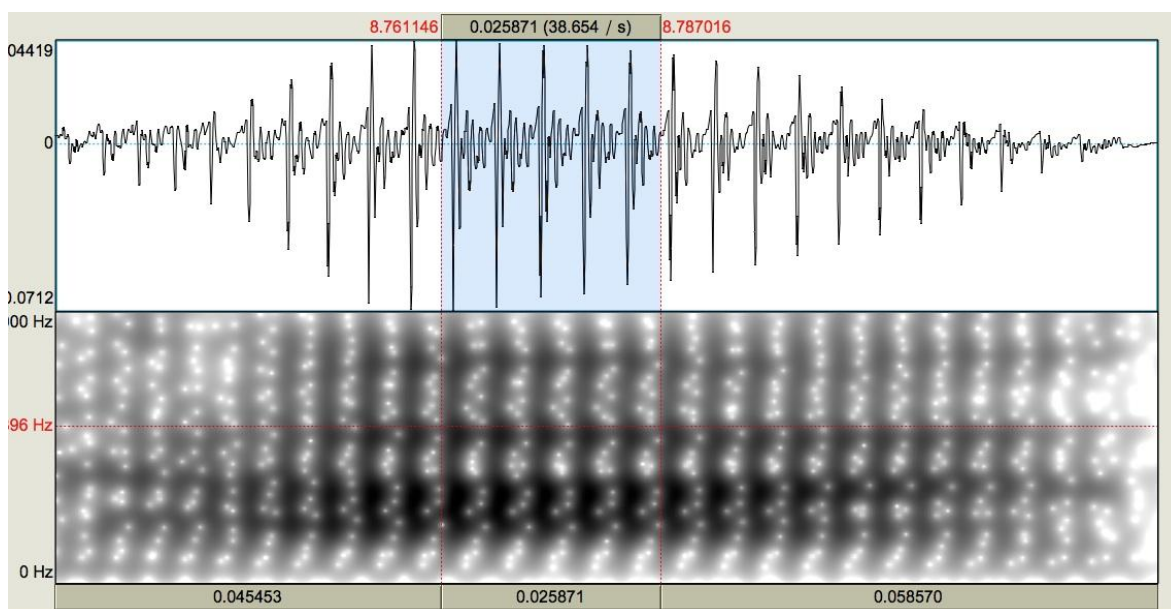


FIGURA 12: CRITÉRIO DE SELEÇÃO DA PORÇÃO DA VOGAL PARA A ANÁLISE DA FONAÇÃO

Durante a seleção da porção da vogal que seria analisada, foi feita também a etiquetagem das vogais em um *TextGrid*, exemplificado na Figura 13:

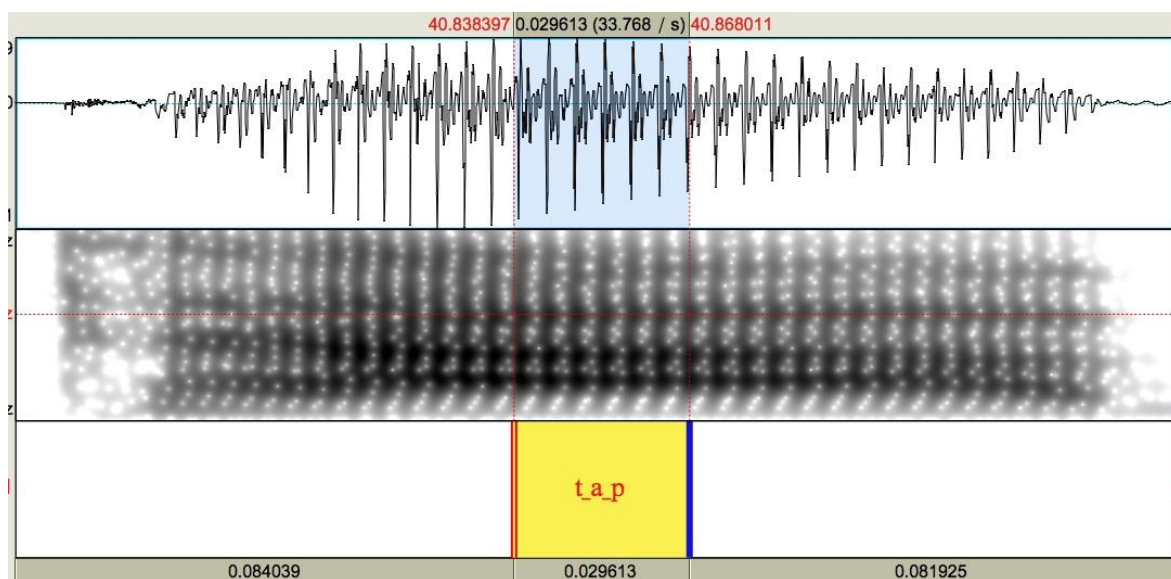


FIGURA 13: ETIQUETAGEM DA PORÇÃO DA VOGAL PARA A ANÁLISE DA FONAÇÃO

Posteriormente à etiquetagem dos dados, as medidas acústicas foram realizadas pelo *software* livre *VoiceSauce* (Shue *et al.*, 2011). Esse programa produz, automaticamente, medidas acústicas a partir de diretórios contendo arquivos *.wav*. Se houver *TextGrids* do Praat para os arquivos de áudio, o programa realiza as medidas apenas nas porções etiquetadas. Pelo programa, podem ser realizadas diversas medidas espectrais e de periodicidade⁶⁸ ao mesmo tempo. Para cada milissegundo, há uma medida de cada parâmetro escolhido; o pesquisador pode optar por obter todas as medidas ou a média entre elas. Depois de feita a análise, o programa gera um arquivo de texto com os resultados. Para a presente análise, foram selecionadas as medidas H1*-H2*, H2*-H4*, H4*-2kHz, 2kHz-5kHz e o PPC, todas expressas em dB. Optou-se por pedir que o programa listasse as médias dessas medidas por vogal.

Das 32 palavras gravadas, 6 foram selecionadas para a realização da análise acústica, sendo todas palavras reais das línguas. A escolha das palavras foi feita com base nas consoantes adjacentes à vogal-alvo (vogal entre consoantes não-vozeadas e entre consoantes vozeadas), no conhecimento das palavras pelos participantes (as palavras mais conhecidas pelos falantes) e pela similaridade entre a palavra em PB e em IN (palavras que tinham a sequência de sons mais próxima possível). Assim, as palavras selecionadas foram: *baba/babble*, *gaba/gabble*, *gago/gaga*, *capa/kappa*, *taco/tacos*, *tapa/tapas*⁶⁹.

5.2.5.2 Medidas de f0 e ELT

As medidas de f0 e ELT foram realizadas nas tarefas de leitura dos textos e da fala semiespontânea. Os valores foram obtidos pelo *Praat*, seguindo os seguintes critérios:

1. Para f0, foram usadas as configurações padrão do programa, a não ser por dois ajustes: (a) a correção das frequências de *pitch* mínima (*floor*) e

⁶⁸ Para todas as opções, ver Shue *et al.* 2011.

⁶⁹ As palavras *tacos* e *tapas* são empréstimos do espanhol; porém, são amplamente utilizadas na Califórnia pela influência espanhola e mexicana na culinária local. Essas palavras foram escolhidas pela grande similaridade que têm com as palavras do português.

máxima (*ceiling*) que eram originalmente de 75 a 600Hz, passando à configuração de 50 a 300Hz para homens, e de 100 a 500Hz para mulheres; (b) o intervalo de tempo de medição, modificado para 10 milissegundos. Depois da extração das medidas, foram selecionadas a média, a extensão e o desvio padrão de cada amostra. No total, foram obtidas dez medidas por falante (três repetições do texto e duas amostras de fala semi-espontânea em cada língua). Todas as medidas de f_0 são expressas em Hz.

2. As medidas do ELT foram obtidas pelo comando *To Ltas (pitch corrected)*, também com correção de *pitch* (50 a 300Hz para homens e 100 a 500Hz para mulheres), com frequência máxima de 8kHz, e largura de banda de 160Hz. As medidas de declínio espectral, em energia, foram geradas pelo comando *Get Slope*, selecionando as faixas de frequência de 0 a 1kHz e de 1 a 5kHz. O resultado desse cálculo é expresso em densidade espectral, que pode ser transformado em energia por uma correção⁷⁰, já que a largura da banda de frequência mais alta é quatro vezes maior que a de frequência mais baixa. No total, foram obtidas dez medidas por falante em cada língua (três repetições do texto e duas amostras de fala semiespontânea). A razão alfa é expressa em dB⁷¹. Além dessa medida, foi realizada outra análise do declínio espectral. Também a partir do *Ltas object* gerado pelo *Praat*, foram calculadas as diferenças entre os picos de amplitude das seguintes bandas de frequência: 0 a 2kHz, 2 a 5kHz e 5 a 8kHz (a diferença entre as duas primeiras faixas e as duas últimas). A primeira diferença será chamada de D0225kHz e a segunda, de D2558kHz. Foram obtidas dez medidas para cada uma das duas diferenças, todas expressas em dB.

⁷⁰ $10 * \log_{10} (4) = 6.0206 \text{ dB}$. Esse número é acrescido ao resultado da densidade espectral, gerando um valor em energia (dB).

⁷¹ Os valores da razão alfa são geralmente negativos, pois representam a queda de energia entre a banda de frequência mais baixa e a mais alta. Neste trabalho, os valores foram transformados em valores positivos, e indicam o quanto a energia é maior na primeira faixa de frequência em relação à segunda.

5.3 EXPERIMENTO DE PERCEPÇÃO

Um dos principais objetivos deste estudo é o de verificar se ouvintes leigos percebem diferenças de qualidade de voz, quando estas ocorrem, na fala de bilíngues. Para acessar esse julgamento, um experimento de percepção foi desenhado a fim de obter tais informações.

O teste de percepção da qualidade de voz de falantes bilíngues foi realizado pelo *software* livre TP – Teste de Percepção (Rauber *et al.*, 2012), na versão 3.1⁷². Foi montada uma tarefa de discriminação, na qual os ouvintes julgaram as vozes dos bilíngues como “iguais” ou “diferentes”.

5.3.1 Os ouvintes

Como ouvintes desse experimento, foram recrutados participantes que não têm experiência com a análise da voz e, portanto, chamados na presente pesquisa de leigos. Os participantes eram professores de LI, com pelo menos um ano de experiência no ensino dessa língua. As condições para que o sujeito participasse do experimento foram: (a) ser brasileiro e falar o PB como língua materna; (b) ter a mesma faixa etária que os participantes do experimento de produção (entre 20 e 39 anos); falar IN como L2 com proficiência; (c) declarar não possuir deficiência auditiva. Por meio de um questionário (ANEXO F), foram levantadas as seguintes informações sobre os participantes: (a) sexo; (b) cidade natal e lugar onde cresceu; (c) idade de aquisição do IN; (d) tempo de estudo de IN; (e) experiência em países onde se fala IN como L1; (f) conhecimento em outra língua estrangeira; (g) autoavaliação de proficiência e sotaque (de 0 a 100, sendo zero “muito sotaque” e 100, “nenhum sotaque”); (h) habilidade musical (tocar instrumento ou cantar). O experimento foi realizado com 14 participantes (sete homens e sete mulheres). Os detalhes sobre os dados biográficos e os relacionados à aquisição do IN estão descritos na Tabela 2 e Tabela 3, respectivamente:

⁷² O TP pode ser obtido no site www.worken.com.br/tp

TABELA 2: IDADE, SEXO, ORIGEM, INÍCIO DE APRENDIZADO E TEMPO DE ESTUDO NO IN DOS OUVINTES

Ouvinte	Idade	Sexo	Origem	Idade início aprendizado	Tempo de estudo no IN
Ouv1	23	F	Curitiba/PR	10	13
Ouv2	39	F	Curitiba/PR	22	17
Ouv3	25	M	Colombo/PR	15	10
Ouv4	29	M	Curitiba/PR	14	15
Ouv5	25	F	Diadema/SP	16	9
Ouv6	22	F	Curitiba/PR	7	15
Ouv7	27	M	Santos/SP	14	13
Ouv8	24	F	Curitiba/PR	12	12
Ouv9	24	F	Curitiba/PR	18	6
Ouv10	24	M	Campo Largo/PR	12	12
Ouv11	23	F	Curitiba/PR	18	5
Ouv12	22	M	Curitiba/PR	8	12
Ouv13	28	M	Curitiba	11	17
Ouv14	24	M	Curitiba	10	14

TABELA 3: EXPERIÊNCIA EM PAÍS DE LI, AUTOAVALIAÇÃO DE PROFICIÊNCIA E SOTAQUE

Ouvinte	Experiência em país de LI (meses)	Autoavaliação Proficiência (%)	Autoavaliação Sotaque estr. (%)
Ouv1	12	90	70
Ouv2	0	70	50
Ouv3	2	90	75
Ouv4	1	60	50
Ouv5	1	60	75
Ouv6	0	100	85
Ouv7	0	70	60
Ouv8	0	85	90
Ouv9	0	60	80
Ouv10	12	90	90
Ouv11	1	70	60
Ouv12	1	80	95
Ouv13	0	80	75
Ouv14	0	70	50

5.3.2 Os estímulos

Os estímulos do teste, baseados nos quais foram feitos os julgamentos, são compostos de trechos da fala semiespontânea produzida pelos participantes do experimento de produção. Foram selecionadas amostras similares semanticamente no PB e no IN – isso foi possível, pois foi instruído ao participante que respondesse

à mesma pergunta em PB e em IN, preferencialmente mantendo conteúdo. Cada *trial* foi composto por um arquivo de áudio em que a amostra em PB foi concatenada à amostra similar em IN, amostras essas do mesmo falante, e apresentadas nessa ordem (PB e depois, IN). O mesmo *trial* foi apresentado duas vezes, totalizando 32 *trials* (duas vezes cada trial dos 16 participantes).

5.3.3 As instruções e a tarefa

Primeiramente, os ouvintes foram instruídos quanto ao uso do *software* e quanto à tarefa que teriam que desempenhar. Foi feita a familiarização dos participantes com o *software* e a tarefa por meio de um *trial* piloto. Esse *trial* piloto foi constituído por uma amostra de fala em PB e em IN de uma pessoa que não participou do experimento de produção, para servir de exemplo aos ouvintes. A tarefa foi explicada aos ouvintes da seguinte forma:

1. Você ouvirá estímulos (*trials*) contendo amostras de fala em PB e em IN, nessa ordem, de uma mesma pessoa. Ao todo, são 16 pessoas diferentes, brasileiros bilíngues falantes de IN como L2. Sua tarefa é analisar e comparar a voz da pessoa quando fala em IN com a voz em PB: não leve em conta desvios sintáticos ou de pronúncia; preste atenção apenas na voz. Você poderá ouvir o mesmo estímulo até cinco vezes antes de se decidir e então deve clicar no botão “igual” ou no botão “diferente” (FIGURA 14):

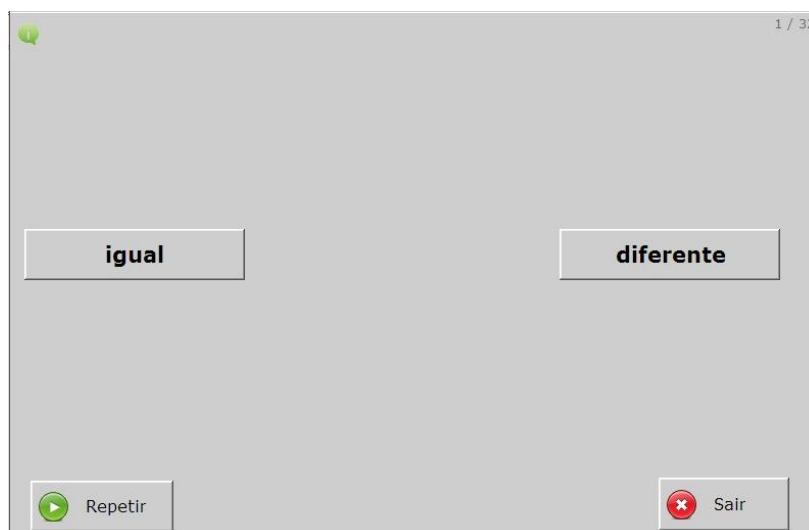


FIGURA 14: TELA DO TP PARA O JULGAMENTO DAS VOZES EM “IGUAL” OU “DIFERENTE”

2. Após clicar em um dos botões “igual” ou “diferente”, você verá uma escala deslizante⁷³ que vai de “pouco” a “muito”. Se você considerou a voz do falante “igual” nas duas línguas, desconsidere a escala deslizante; apenas utilize-a para quantificar a diferença caso tenha clicado em “diferente”. Você deve quantificar a diferença entre “pouco” e “muito” (FIGURA 15).

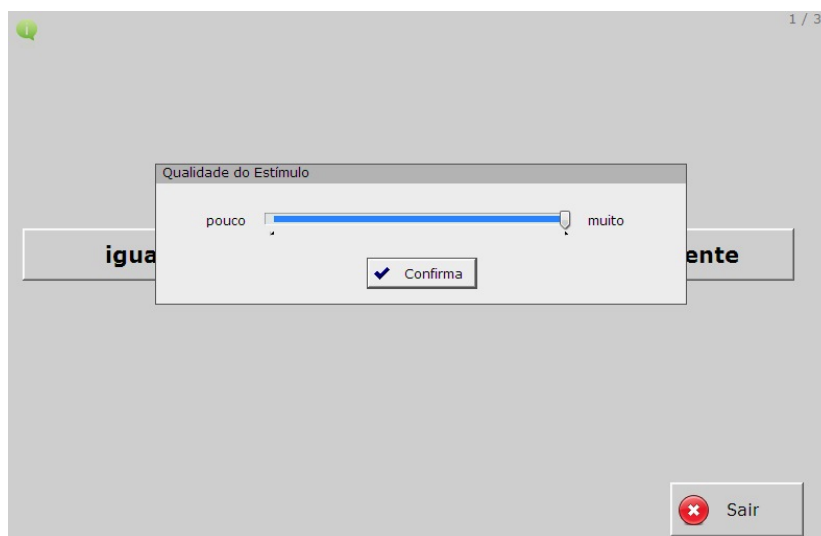


FIGURA 15: TELA DO TP COM A ESCALA DESLIZANTE PARA A QUANTIFICAÇÃO DA DIFERENÇA ENTRE AS VOZES

3. Caso você julgue que a voz seja diferente no IN em relação ao PB, você terá que dizer qual é essa diferença: você poderá descrever a diferença usando a palavra ou expressão que desejar⁷⁴.

A partir dos julgamentos, foram geradas planilhas com os resultados da discriminação entre “igual” e “diferente” e com a quantificação das diferenças. Se o ouvinte considerou que o falante soava igual em IN e PB, o valor atribuído ao julgamento foi de zero. Se o ouvinte considerou que o falante soava diferente em IN em comparação ao PB e quantificou essa diferença na escala deslizante, foram

⁷³ Essa escala deslizante gera um valor de 0 a 100.

⁷⁴ Não foram sugeridos termos que descrevessem a voz do falante. Esperava-se que o ouvinte utilizasse adjetivos que descrevessem tanto a voz propriamente dita, como características emocionais ou de personalidade do falante. Foi pedido ao ouvinte que falasse o descritor, e a pesquisadora anotava essa palavra ou expressão, já que não havia a possibilidade de escrevê-la no *software*.

gerados valores de 1 a 100, sendo 1 considerado muito pouca diferença e 100, muita diferença.

5.4 DESCRIÇÃO E ANÁLISE DOS DADOS

Nesta seção, são apresentados os dados brutos obtidos a partir dos experimentos de produção e percepção, bem como os procedimentos de análise estatística empregados a fim de testar as hipóteses apresentadas no capítulo 1.

A partir do experimento de produção, foram gerados dados nos seguintes totais:

- a. 2880 médias relacionadas às medidas de curto termo – H1*-H2*, H2*-H4*, H4*-2kHz, 2kHz-5kHz e PPC (cinco medidas X seis vogais X duas línguas X três repetições X 16 participantes).
- b. 288 médias relacionadas às medidas de declínio espectral do ELT - razão alfa, D0225kHz e D2558kHz – na tarefa de leitura de texto (três medidas X duas línguas X três repetições X 16 participantes); e 192 medidas realizadas na fala semiespontânea (três medidas X duas línguas X duas repetições X 16 participantes).
- c. 192 medidas relacionadas à média e extensão da f0 na tarefa de leitura de texto em PB e em IN (duas medidas X duas línguas X três repetições X 16 participantes) e 128 medidas obtidas nas amostras de fala semi-espontânea em PB e em IN (duas medidas X duas línguas X duas repetições X 16 participantes).

Com relação ao experimento de percepção, foram gerados 448 dados (16 trials X 2 repetições X 14 ouvintes).

Quanto à análise estatística, primeiramente foram calculadas a média e o desvio-padrão dos dados do experimento de produção, bem como foram inspecionadas a assimetria e a curtose dos mesmos. Após verificar que a maioria dos dados não apresentava distribuição normal, foram utilizados testes não-paramétricos para a análise estatística inferencial. A escolha dos testes foi feita com base na natureza das perguntas de pesquisa e hipóteses, a saber: *design*

intrassujeito para verificar diferenças entre as duas línguas (teste Wilcoxon) para as perguntas de pesquisa 1, 2 e 3; *design* intersujeito para verificar diferenças entre homens e mulheres (teste Mann-Whitney) para a pergunta 4; correlação entre a qualidade de voz e variáveis relacionadas à aquisição do IN (teste Pearson) para a pergunta 5. Para que os resultados fossem estatisticamente significativos, foi considerado o valor de $p \leq 0,05$.

Para responder às perguntas de pesquisa 6 e 7 sobre a percepção da qualidade de voz, foi reportada a estatística descritiva dos dados: média e desvio-padrão dos julgamentos dos 14 ouvintes em relação à produção dos 16 falantes⁷⁵. Além disso, foi rodado o teste Pearson para verificar a correlação entre os julgamentos a fim de verificar a concordância entre os ouvintes. Baseado nos resultados desse teste, foi criado um subgrupo dos dez ouvintes que obtiveram maior número correlação uns com os outros nos julgamentos. Ademais, os falantes foram agrupados (grupo 1 e grupo 2) conforme o julgamento desse subgrupo de ouvintes. Para verificar a relação entre os resultados dos testes de produção e de percepção, foi rodado o teste Kruskal-Wallis para verificar diferenças entre os três grupos de falantes (pouco sotaque, médio sotaque e muito sotaque estrangeiro) utilizando as medidas de declínio espectral e de f0.

5.5 CONCLUSÃO DO CAPÍTULO

Neste capítulo, foram relatadas as características metodológicas dos experimentos de produção e de percepção da presente pesquisa. Foram descritos os participantes, os materiais e métodos para a coleta e análise dos dados. Além disso, foram descritos os procedimentos de análise dos dados. A partir dos resultados dessas duas partes da presente pesquisa, será possível testar as hipóteses descritas na introdução e relacionar a produção da voz dos bilíngues e a percepção da qualidade de voz por ouvintes leigos. No capítulo seguinte, serão apresentados os resultados dos dois experimentos realizados para esta pesquisa.

⁷⁵ Ao reportar os resultados do experimento de percepção, será utilizada a palavra “falantes” para designar os participantes do experimento de produção.

6 RESULTADOS

A partir das medidas geradas pelos programas *VoiceSauce* e *Praat*, foi utilizado o programa de análise estatística *Statistical Package for Social Sciences* (IBM SPSS, versão 22), para tratar os dados estatisticamente. Primeiramente, são apresentados os resultados da análise estatística do experimento de produção. Em um segundo momento, são reportados os resultados do experimento de percepção. Para cada experimento, são apresentados os resultados de uma análise estatística descritiva de cada pergunta de pesquisa e hipótese e, em seguida, são reportados os resultados da análise estatística inferencial.

6.1 EXPERIMENTO DE PRODUÇÃO

A partir do experimento de produção, foi gerado um grande número de medidas acústicas que descrevem a qualidade de voz dos participantes brasileiros bilíngues. Por isso, foi realizado o agrupamento dos dados produzidos pelos participantes quando não havia diferença estatisticamente significativa entre as repetições. Depois disso, são apresentadas a média e o desvio-padrão de todas as medidas (espectrais de curto termo e longo termo, e f0), com os valores de homens e mulheres separadamente. Por fim, são testadas as hipóteses advindas das perguntas de pesquisa pela estatística inferencial, seguidas da discussão dos resultados significativos e relação com outros estudos.

6.1.1 Agrupamento dos dados e estatística descritiva

Para que fosse possível fazer generalizações sobre as produções em PB e em IN, foram realizados testes para verificar a existência de diferenças estatisticamente significativas entre as repetições e entre as palavras nas quais as medidas foram realizadas. Caso não houvesse tais diferenças, os dados seriam agrupados.

6.1.1.2 Medidas espectrais de curto termo

Para cada uma das cinco medidas de curto-termo, havia 36 médias para cada falante: seis palavras X três repetições X duas línguas. Dessas 36 médias, 18 eram medidas de vogais que estavam entre as consoantes vozeadas (/bab/, /gab/ e /gag/) e as outras três entre as não-vozeadas (/kap/, /tak/ e /tap/). Primeiramente, foi rodado o teste Friedman para verificar se havia diferença estatística entre as três repetições da mesma palavra, em cada língua (exemplo: bab1 X bab2 X bab3 do PB). Os resultados mostraram não haver diferenças estatisticamente significativas entre as três repetições de cada palavra ($p > ,05$). Dessa forma, foi calculada a média entre as três repetições de cada palavra, gerando 12 médias para cada falante (seis palavras X duas línguas). Posteriormente, foi rodado o mesmo teste para verificar se havia diferença estatística entre as medidas realizadas nas vogais em função do ponto de articulação das consoantes adjacentes (exemplo: bab X gab X gag do PB). Novamente, os resultados mostraram não haver diferença estatisticamente significativa entre as medidas das vogais em função do ponto de articulação ($p > ,05$). Assim, foi calculada a média entre o ambiente vozeado e o não-vozeado, resultando em 4 médias (vozeadas ou não-vozeadas X duas línguas). Por fim, foi rodado o teste Wilcoxon para verificar se havia diferença estatística entre as medidas em função do vozeamento das consoantes adjacentes (vozeadas X não-vozeadas do PB). Houve diferença estatisticamente significativa na medida PPC em duas comparações: entre os ambientes vozeado e não-vozeado do IN ($Z = -2,585$, $p = ,01$) e vozeado e o não-vozeado do PB ($Z = -2,999$, $p = ,003$). Os resultados das outras medidas mostraram não haver diferença estatisticamente significativa entre as vogais em função do vozeamento das consoantes adjacentes. Portanto, os dados foram agrupados de maneira que, para cada medida acústica, havia uma média para o PB e outra para o IN, a não ser pela medida de PPC.

Assim sendo, as médias e os DP das medidas acústicas de curto termo, a não ser pelo PPC, são apresentadas na Tabela 4, separadas por sexo do participante⁷⁶:

TABELA 4: MÉDIA E DP DAS MEDIDAS DE CURTO TERMO PARA OS PARTICIPANTES DO SEXO MASCULINO (N=6) E DO SEXO FEMININO (N=10), EM DECIBÉIS (dB)

Medida acústica	Homens – PB Média (DP)	Homens – IN Média (DP)	Mulheres – PB Média (DP)	Mulheres – IN Média (DP)
H1*-H2*	-3,52 (3,12)	-3,97 (3,96)	3,18 (2,14)	3,44 (3,19)
H2*-H4*	4,64 (2,24)	4,96 (2,57)	6,40 (3,10)	6,75 (2,93)
H4*-2kHz	5,44 (3,00)	4,61 (3,81)	5,34 (5,23)	4,94 (5,71)
2kHz-5kHz	22,91 (2,97)	22,1 (3,55)	14,02 (5,08)	12,63 (5,12)

Como já reportado na literatura, os valores de H1*-H2* são mais baixos para os homens, indicando a produção da voz crepitante, do que para as mulheres, que produzem uma voz mais soprosa. Com relação à medida H2*-H4*, as mulheres apresentaram valores mais altos, também sinalizando a presença de mais sopro na voz em relação à dos homens. Por fim, quanto à medida 2kHz-5kHz, as mulheres apresentaram médias mais baixas que as dos homens, justamente pela presença de maior ruído no sinal nessa faixa de frequência.

As médias do PPC, por serem diferentes em função do vozeamento nas duas línguas, são apresentadas na Tabela 5:

Tabela 5: MÉDIA E DP DAS DO PPC PARA OS PARTICIPANTES DO SEXO MASCULINO (N=6) E DO SEXO FEMININO (N=10), EM Db

Medida acústica	Homens – PB Média (DP)	Homens – IN Média (DP)	Mulheres – PB Média (DP)	Mulheres – IN Média (DP)
PPC – voz.	25,68 (1,33)	26,04 (1,78)	25,62 (2,36)	24,81 (3,25)
PPC – não-voz.	23,79 (2,33)	24,24 (2,47)	24,29 (3,86)	22,90 (3,52)

A medida do PPC indica a regularidade dos picos harmônicos do sinal acústico, de forma que se a estrutura harmônica for bem definida, o valor do PPC é alto, ao passo que se for baixo, indica presença de ruído no sinal. Para o ambiente

⁷⁶ Os resultados da estatística descritiva dos dados são apresentados separando homens e mulheres por conta da conhecida diferença na produção da voz entre os dois sexos. Os resultados da análise inferencial intrassujeito, contudo, são apresentados sem essa separação.

vozeado, os homens apresentam médias do PPC um pouco maiores que as mulheres, assim como no ambiente não-vozeado do inglês. Por outro lado, no ambiente não-vozeado do PB, a média das mulheres foi superior a dos homens.

6.1.1.3 Medidas espectrais de longo termo

Assim como nas medidas de curto termo, foram realizados testes estatísticos para agrupar os dados que não apresentassem diferenças significativas. Primeiramente, com o objetivo de verificar se havia diferenças intrassujeito entre as três repetições da leitura do texto quanto às medidas de declínio espectral do ELT, foi rodado o teste Friedman para cada medida em cada língua. Os resultados mostraram não haver diferenças estatisticamente significativas entre as três repetições em cada língua ($p > 0,05$). Portanto, a partir das médias das três repetições, foi calculada uma média referente a cada uma das três medidas para cada língua (TABELA 6):

TABELA 6: MÉDIA E DP DAS MEDIDAS DE DECLÍNIO ESPECTRAL DE ELT DA LEITURA DE TEXTO DOS PARTICIPANTES DO SEXO MASCULINO (N=6) E DO FEMININO (N=10), EM dB.

Medida de ELT	Homens – PB Média (DP)	Homens – IN Média (DP)	Mulheres – PB Média (DP)	Mulheres – IN Média (DP)
Razão alfa	10,44 (1,38)	9,82 (2,09)	10,03 (2,44)	9,25 (2,81)
D0225kHz	19,10 (2,84)	17,60 (1,76)	19,74 (3,5)	16,76 (3,56)
D2558kHz	16,95 (2,45)	19,45 (4,04)	11,46 (5,07)	16,39 (4,76)

Na tarefa de leitura de texto, as médias da razão alfa mostram declínio espectral parecido entre homens e mulheres, tendo as vozes em IN médias um pouco mais baixas que as do PB. Esses números parecem indicar vozes mais soprosas em PB que em IN, além de poder indicar que as vozes em IN foram produzidas com maior intensidade. Com relação à medida D0225kHz, as médias das produções de homens e mulheres foram semelhantes. Da mesma forma, o declínio espectral é um pouco maior em PB que em IN, confirmando a tendência de vozes mais soprosas naquela língua. Contudo, a medida D2558kHz teve médias mais altas para homens do que para mulheres, indicando voz crepitante para eles

na região mais alta de frequência do espectro. Pode-se notar também que tanto homens quanto mulheres produziram voz mais crepitante em IN que em PB.

Com relação à fala semiespontânea, foi realizado o teste Wilcoxon para verificar possíveis diferenças entre as duas amostras produzidas em cada língua. Novamente, não houve diferenças estatisticamente significativas entre as duas amostras ($p>0,05$), possibilitando que fosse calculada uma média para cada medida de declínio espectral em cada língua (TABELA 7):

TABELA 7: MÉDIA E DP DAS MEDIDAS DE DECLÍNIO ESPECTRAL DE ELT DA FALA SEMI-ESPONTÂNEA DOS PARTICIPANTES DO SEXO MASCULINO (N=6) E DO FEMININO (N=10), EM dB.

Medida de ELT	Homens – PB Média (DP)	Homens – IN Média (DP)	Mulheres – PB Média (DP)	Mulheres – IN Média (DP)
Razão alfa	9,59 (1,49)	9,75 (2,03)	8,07 (2,28)	7,60 (2,79)
D0225kHz	17,72 (2,96)	17,62 (2,03)	15,90 (2,77)	14,95 (3,00)
D2558kHz	18,61 (2,74)	19,80 (2,95)	15,91 (4,09)	17,16 (4,74)

Os resultados das medidas espectrais da fala semiespontânea foram mais altas para os homens do que para as mulheres. Com relação à razão alfa, os homens produziram valores bastante similares em PB e em IN. As mulheres, por sua vez, obtiveram média mais alta em PB que em IN, indicando declínio espectral mais acentuado e, assim, produzindo vozes mais soprosas na sua L1. A respeito da diferença entre as médias dos homens e mulheres, pode ser que elas tenham produzido vozes com mais intensidade que eles, resultando em um declínio espectral mais atenuado pela maior concentração de energia na faixa mais alta de frequência. Assim também aconteceu com a D0225kHz entre as línguas e entre os sexos. Com referência à medida D2558kHz, os homens apresentaram maior declínio espectral, indicando assim a produção de vozes mais crepitantes que as das mulheres.

6.1.1.4 Medidas de f0

Assim como para as medidas espectrais, foram rodados testes estatísticos para averiguar possíveis diferenças entre as repetições da leitura do texto e da fala semiespontânea. Primeiramente, a partir do resultado do teste Friedman, houve

diferença estatisticamente significativa entre as médias de f0 das três repetições da leitura do texto em inglês ($\chi^2(2)=6$, $p=0,05$). Por conta dessa diferença, foi rodado o teste Wilcoxon entre os pares (TX-IN-1) e (TX-IN-2), (TX-IN-2) e (TX-IN-3), e (TX-IN-1) e (TX-IN-3), utilizando a correção de Bonferroni (alfa: $0,05/3=0,0017$). Contudo, aplicada a correção de Bonferroni, não foi verificada diferença estatisticamente significativa entre as repetições⁷⁷. Além disso, não houve diferenças estatisticamente significativas nas comparações entre as repetições do PB. Portanto, foram calculadas as médias das medidas de f0 obtidas nas três repetições da leitura dos textos em cada língua (TABELA 8):

TABELA 8: MÉDIA E DP DAS MEDIDAS DE F0 DA LEITURA DE TEXTOS EM PB E IN POR PARTICIPANTES HOMENS (N=6) E MULHERES (N=10), EM Hz

Medida de f0	Homens – PB Média (DP)	Homens – IN Média (DP)	Mulheres – PB Média (DP)	Mulheres – IN Média (DP)
Média	114,50 (10,39)	110,98 (9,05)	208,63 (24,04)	205,03 (25,78)
Extensão	223,91 (14,08)	229,60 (23,15)	372,82 (14,64)	380,82 (8,42)

Na tarefa de leitura de texto, as diferenças entre homens e mulheres no que concerne à f0 segue o padrão já amplamente descrito na literatura: os homens têm médias de f0 mais baixas que as mulheres. Além disso, a extensão do *pitch* das mulheres é maior que a extensão que os homens produzem.

Para verificar se as medidas relacionadas à média e extensão da f0 da fala semiespontânea apresentavam diferenças entre si, foi rodado o teste Wilcoxon entre as duas amostras de cada língua. Novamente, não foram encontradas diferenças estatisticamente significativas entre as amostras de fala semiespontânea de cada língua ($p>0,05$). Dessa forma, foram calculadas as médias de cada medida em cada língua (TABELA 9):

⁷⁷ A diferença revelada pelo teste Friedman foi detectada pelo teste de Wilcoxon entre as repetições 1 e 3 do texto em IN ($Z=-2,223$, $p=0,026$). Assim, o valor de p foi maior que a estipulada pela correção de Bonferroni de 0,017 e, portanto, considerou-se não haver diferenças estatisticamente significativas entre essas duas repetições.

TABELA 9: MÉDIA E DP DAS MEDIDAS DE F0 DA FALA SEMI-ESPONTÂNEA EM PB E IN POR PARTICIPANTES HOMENS (N=6) E MULHERES (N=10), EM Hz

Medida de f0	Homens – PB Média (DP)	Homens – IN Média (DP)	Mulheres – PB Média (DP)	Mulheres – IN Média (DP)
Média	112,90 (9,85)	111,15 (10,64)	205,59 (20,13)	200,32 (24,13)
Extensão	230,59 (20,78)	236,52 (8,93)	363,43 (37,59)	366,05 (29,69)

Por fim, na apresentação da média e extensão da f0 na fala semiespontânea, foram verificadas as mesmas tendências da produção da tarefa de leitura de texto: média e extensão mais altas para as mulheres.

6.1.2 Testagem das hipóteses

6.1.2.1 Hipótese 1

Para testar a hipótese 1, a respeito da diferença nas medidas espectrais de curto-termo na comparação entre as produções do PB e do IN, foi rodado o teste Wilcoxon. O resultado do teste mostrou não haver diferença estatisticamente significativa nas medidas espectrais entre a produção do PB e a do IN dos brasileiros bilíngues que participaram do experimento.

TABELA 10: RESULTADOS DO TESTE WILCOXON ENTRE PB E IN PARA MEDIDAS DE CURTO TERMO

	H1-H2 PB x IN	H2-H4 PB x IN	H4-2K PB x IN	2K-5K PB x IN	PPC_Voz PB x IN	PPC_NVoz PB x IN
Z	-0,103	-0,672	-0,052	-1,913	-1,862	-1,448
Significância (p)	,918	,501	,959	,056	,063	,148

Segundo a Tabela 10, a medida 2kHz-5kHz teve apenas uma diferença estatística marginalmente significativa, com a média do IN sendo um pouco mais baixa que a do PB, assim indicando presença de mais ruído nessa faixa de frequência no IN que no PB.

Para verificar se havia diferenças entre as medidas do PB e do IN no grupo dos homens e no das mulheres separadamente, o teste Wilcoxon foi novamente rodado para cada sexo. Foi encontrada diferença estatisticamente significativa em uma das medidas no grupo das mulheres – PPC em ambiente não vozeado (Z=-

2,090, $p=0,037$). O PPC indica maior sopro na voz, segundo Hillebrand (1986). Essa sopro pode ser justificada pela influência do ambiente não-vozeado, já que no inglês, as oclusivas não-vozeadas são aspiradas. Assim, as brasileiras parecem estar estendendo a aspiração para a fonação da vogal adjacente; isso pode indicar que elas ainda não tenham adquirido o faseamento adequado da aspiração das consoantes não-vozeadas do inglês. Não houve diferenças estatisticamente significativas no grupo dos homens.

Portanto, a hipótese nula foi mantida. Por não haver no PB nem no IN fonação contrastiva, era esperado que não houvesse diferenças entre a produção do PB e do IN. Como não há outros estudos que façam a comparação dessas medidas espectrais entre línguas em amostras do curto termo, não se pode relacionar esses resultados a outras pesquisas na área.

6.1.2.2 Hipóteses 2.1 e 2.2

Da mesma forma, para testar a hipótese 2.1, sobre a diferença nas características espectrais de longo termo na comparação entre a produção do PB e do IN, foi utilizado o teste Wilcoxon, cujo resultado é apresentado na Tabela 11:

TABELA 11: RESULTADOS DO TESTE WILCOXON ENTRE PB E IN PARA MEDIDAS ESPECTRAIS DE LONGO TERMO

	TX	SE
	PB X IN	PB X IN
Razão Alfa	Z=-2,534, $p=,011$	Z=-1,034 $p=,301$
D0225kHz	Z=-3,516, $p=,000$	Z=-1,500, $p=,134$
D2558kHz	Z=-3,361, $p=,001$	Z=-2,327, $p=,020$

Os resultados da Tabela 11 mostram que há diferenças estatisticamente significativas na comparação entre a produção em IN e em PB na tarefa de leitura de texto. A razão alfa e a D0225kHz tiveram médias mais altas em PB. Segundo Kitzing (1986), a noção psicoacústica de sonoridade na qualidade de voz é refletida em um declínio espectral menos acentuado. Portanto, a partir dos resultados dessas duas medidas, pode-se afirmar que as vozes dos participantes tiveram maior intensidade no IN que no PB. Keating e Garellek (2015) afirmam que a voz

crepitante tem harmônicos mais fortes na região entre 3 e 5kHz, justamente o que aconteceu na produção em IN descrita pelo declínio espectral menos acentuado nessa língua. Assim, de maneira geral, os brasileiros produziram uma voz mais soprosa no PB que no IN. A medida D2558kHz também revelou essa tendência, pois as médias foram mais altas em IN que em PB, indicando maior presença de voz crepitante nessa língua. Segundo Hammarberg *et al.* (1980), a voz soprosa está relacionada com um maior declínio espectral entre as faixas de 0 a 2kHz e 2 a 5kHz, enquanto que a voz crepitante, está relacionada com um maior declínio espectral entre as faixas 2 a 5kHz e 5 a 8kHz.

Com relação à fala semiespontânea, houve diferença estatisticamente significativa na medida D2558kHz entre o PB e o IN, indicando a presença de mais crepitação na voz no IN.

Ao analisar o grupo dos homens separado do grupo das mulheres, constatou-se que para eles, a única diferença estatisticamente significativa foi a da medida D0225kHz na tarefa de leitura de texto. Esses participantes produziram vozes menos soprosas e mais tensas no IN que no PB. O grupo das mulheres apresentou diferenças estatisticamente significativas nas três medidas na tarefa de leitura de texto, produzindo mais soprosidade no PB e crepitação no IN.

Os presentes resultados estão em consonância com os de Ng *et al* (2012), que compararam as características espectrais de bilíngues cantonês/IN na tarefa de leitura de texto. Nesse estudo, as participantes mulheres apresentaram declínio espectral mais baixo que os homens (médias mais altas). Além disso, todos os participantes apresentaram declínio espectral mais baixo em cantonês, sua língua materna, do que em IN. No estudo de Todaka (1995), da mesma forma, as mulheres apresentaram configuração laríngea mais relaxada e constante abertura glotal em japonês que em IN.

Com o objetivo de testar a hipótese 2.2, e verificar possíveis diferenças entre as tarefas, foi rodado novamente o teste Wilcoxon para comparar as medidas espectrais das produções dos participantes bilíngues: TX-IN versus SE-IN; TX-PB versus SE-PB, conforme a tabela 12:

TABELA 12: RESULTADOS DO TESTE WILCOXON PARA MEDIDAS DE ELT ENTRE TAREFAS LEITURA DE TEXTO (TX) E FALA SEMIESPONTÂNEA (SE)

	TX-IN x SE-IN	TX-PB x SE-PB
Razão alfa	Z=-2,327, p=,020	Z=-2,947, p=,003
D0225kHz	Z=-1,603, p=,109	Z=-3,206, p=,001
D2558kHz	Z=-0,621, p=,535	Z=-3,154, p=,002

De acordo com a Tabela 12, nas comparações entre tarefas do IN, houve diferença estatisticamente significativa apenas na razão alfa, sendo a média mais baixa na fala semiespontânea que na leitura de texto. Isso indica que os participantes produziram vozes mais soprosas na leitura de texto em relação à fala semiespontânea. Já no PB, houve diferenças estatisticamente significativas nas três medidas na comparação entre os tipos de tarefas. O declínio espectral foi mais alto na leitura de texto em comparação com a fala semiespontânea nas medidas razão alfa e D0225kHz; na medida D2558kHz, o declínio é maior na fala semiespontânea que na leitura do texto. Esses resultados indicam maior sopro na leitura de texto no PB e maior crepitação na fala semiespontânea no IN.

Camargo *et al.* (2013) também encontraram diferenças estatisticamente significativas entre a tarefa de leitura de texto em comparação com a fala semiespontânea nas medidas de declínio espectral. Esses autores reportaram valores mais altos de declínio espectral na fala semiespontânea nas três línguas do participante.

Além disso, o fato de no IN haver menos diferenças entre as tarefas está em consonância com as conclusões dos estudos de Bruyninckx *et al.* (1991; 1994), que mostraram haver mais consistência/menos variabilidade na língua não-dominante. Portanto, nossos resultados parecem indicar que os participantes bilíngues têm maior variabilidade quando falam o PB do que quando produzem o IN.

Dividindo os participantes entre homens e mulheres, tem-se que o grupo dos homens não apresenta nenhuma diferença estatisticamente significativa, enquanto que no grupo das mulheres, houve diferença nas mesmas medidas do grupo como um todo.

6.1.2.3 Hipóteses 3.1 e 3.2

Com relação às medidas de f_0 entre as línguas, foram obtidos os resultados a partir do teste Wilcoxon mostrados na Tabela 13:

TABELA 13: RESULTADOS DO TESTE WILCOXON ENTRE PB E IN PARA MEDIDAS DE f_0

Medida de f_0	TX PB X IN	SE PB X IN
Média	Z=-1,965, p=,049	Z=-1,655, p=,098
Extensão	Z=-1,965, p=,049	Z=-0,517, p=,605

Segundo a Tabela 13, na tarefa de leitura de texto, as duas medidas de f_0 tiveram diferença estatisticamente significativa entre o PB e o IN. A média de f_0 foi mais baixa no IN que no PB, enquanto que a extensão foi mais alta. Isso indica que, por mais que os brasileiros tenham um *pitch* mais baixo na leitura em IN, eles também apresentam maior variação na f_0 nessa língua. Por outro lado, não houve diferença estatisticamente significativa entre o PB e o IN na fala semiespontânea.

Outros estudos também verificaram valores de f_0 mais altos para as línguas maternas do que para as línguas estrangeiras. Todaka (1995) constatou que suas participantes produziam f_0 mais altas em japonês que em inglês; as bilíngues russo/IN de Altenberg e Ferrand (2006) produziram f_0 mais alta em russo que em inglês. Por outro lado, Ng *et al.* (2012), ao comparar as médias de f_0 em inglês e cantonês, verificaram que as participantes produziam f_0 mais alto em inglês, sua L2.

A respeito das medidas de f_0 na comparação as tarefas desempenhada pelo participante, obteve-se os seguintes resultados (TABELA 14):

TABELA 14: RESULTADOS DO TESTE WILCOXON PARA MEDIDAS DE f_0 DO ENTRE AS TAREFAS LEITURA DE TEXTO (TX) E DE FALA SEMI-ESPONTÂNEA (SE)

Medida de f_0	TX-IN x SE-IN	TX-PB x SE-PB
Média	Z=-1,344, p=,179	Z=-1,603, p=,109
Extensão	Z=-0,827, p=,408	Z=-0,078, p=,938

A Tabela 14 mostra que os participantes não apresentam diferença estatisticamente significativa entre as medidas de f0 produzidas em PB e as do IN na comparação entre tarefas. Embora não haja diferença estatisticamente significativa entre as tarefas, pode-se constatar pelas medidas que as médias produzidas na fala semiespontânea foram mais baixas que as na leitura de texto.

6.1.2.4 Hipótese 4

Essa hipótese testa diferenças entre as medidas espectrais de curto e de longo termo, bem como as de f0 entre homens e mulheres. O teste utilizado foi o Mann Whitney. São apresentados apenas os resultados estatisticamente significativos:

TABELA 15: RESULTADOS DO TESTE MANN WHITNEY ENTRE HOMENS E MULHERES PARA AS MEDIDAS ESPECTRAIS DE CURTO E LONGO TERMO

Medidas espectrais	H1*-H2* IN	H1*-H2* PB	2k-5kHz IN	2k-5kHz PB	D2558 TX PB
Z	-3,145	-3,254	-2,929	-2,929	-2,278
Significância (p)	,002	,001	,003	,003	,023

De acordo com a tabela 15, as vozes produzidas pelos homens e mulheres que participaram deste estudo apresentaram diferenças estatisticamente significativas em duas das quatro medidas de curto-termo nas duas línguas. Em ambos PB e IN, os homens tiveram médias mais baixas de H1*-H2* que as mulheres. Eles produziram as vogais com voz crepitante (médias com valores negativos) e as mulheres com a voz mais soprosa (médias com valores positivos). Essas diferenças já foram reportadas pela literatura: Pépiot (2014) reportou que as mulheres participantes de seu estudo, tanto as francesas quanto as estadunidenses, apresentaram valores de H1*-H2* mais altos que os homens, sendo que os homens estadunidenses produziram H1*-H2* ainda mais baixos que os homens franceses. Ademais, Podesva (2013) relata que o padrão de voz crepitante está associado à masculinidade, já que eles geralmente apresentam *pitch* mais baixo que as mulheres, e um mais baixo favorece a produção da voz crepitante.

Com relação à medida 2kHz-5kHz na leitura de texto no PB, as mulheres tiveram médias mais baixas que os homens, indicando maior presença de energia nessa faixa de frequência. Essa diferença pode explicar uma maior soproidade na fala feminina. Para Mendoza *et al.* (1996), as mulheres apresentam maior nível de energia na região espectral de F3, o que as faz ter vozes perceptualmente mais soprosas. Garellek *et al.* (2013) afirmam que, quando o nível de ruído é alto nessa faixa de frequência, o declínio espectral é menor, presumidamente pela geração de ruído aspirado na abertura glotal.

A respeito da medida D2558kHz, os homens tiveram declínio espectral mais acentuado, indicando maior incidência de voz crepitante na leitura do texto em PB. Por conseguinte, as mulheres apresentaram um declínio mais suave, indicando a presença de ruído nessa faixa mais alta de frequência.

No que concerne às diferenças nas medidas de f0 entre homens e mulheres, foi rodado o teste Mann Whitney, a partir do qual foram obtidos os resultados mostrados nas Tabelas 16 e 17:

TABELA 16: RESULTADOS DO TESTE MANN WHITNEY ENTRE HOMENS E MULHERES PARA AS MEDIDAS DE F0 NA LEITURA DE TEXTO (TX)

Medidas de f0 TX	Média f0 TX IN	Média f0 TX PB	Extensão f0 TX IN	Extensão f0 TX PB
Z	-3,254	-3,254	-3,254	-3,256
Significância (p)	,001	,001	,001	,001

TABELA 17: RESULTADOS DO TESTE MANN WHITNEY ENTRE HOMENS E MULHERES PARA AS MEDIDAS DE F0 NA FALA SEMIESPONTÂNEA (SE)

Medidas de f0 SE	Média f0 SE IN	Média f0 SE PB	Extensão f0 SE IN	Extensão f0 SE PB
Z	-3,254	-3,254	-3,254	-3,254
Significância (p)	,001	,001	,001	,001

De acordo com as Tabelas 16 e 17, como já era esperado, houve diferença estatisticamente significativa entre a produção da voz de homens e mulheres em todas as medidas de f0, nas duas línguas e nas duas tarefas. Segundo Kreiman e Sidtis (2011), por conta de diferenças fisiológicas (principalmente o tamanho da

faringe e da laringe) a f_0 de homens e mulheres diferem-se por aproximadamente uma oitava, tendo eles uma média de 130Hz e elas, 220Hz.

6.1.2.5 Hipótese 5

A hipótese 5 diz respeito à possível correlação entre o tempo de experiência em países de língua inglesa e idade de aquisição com a produção da voz em IN. Além disso, foi testada a existência de diferenças nas emissões do IN entre os grupos separados pelo julgamento de grau de sotaque estrangeiro por falantes nativos dessa língua.

Para verificar a correlação entre o tempo de experiência em países de LI e idade de aquisição com as medidas acústicas obtidas a partir da fala em IN, foi utilizado o teste Pearson, cujos resultados estão na Tabela 18:

TABELA 18: RESULTADOS DO TESTE PEARSON ENTRE A IDADE DE AQUISIÇÃO E AS MEDIDAS ACÚSTICAS

Medidas em IN	H1*-H2*	Média f_0 TX	Extensão f_0 TX	Média f_0 SE	Extensão f_0 SE
Idade de aquisição	$r = ,543$ $p = ,030$	$r = ,622$ $p = ,010$	$r = ,562$ $p = ,023$	$r = ,625$ $p = ,010$	$r = ,603$ $p = ,013$

Somente foram reportados os resultados com significância estatística. Assim, segundo a Tabela 18, não houve correlação entre o tempo de experiência em países de língua inglesa e a qualidade de voz em IN dos participantes desta pesquisa. Por outro lado, houve correlação entre a idade de aquisição do IN e a medida espectral de curto termo H1*-H2* e todas as medidas de f_0 .

Com relação à medida H1*-H2*, quanto mais cedo tenha ocorrido o início da aquisição do IN pelo participante, mais baixo foi o valor de H1*-H2*. Como os participantes desta pesquisa produziram valores de f_0 mais baixos em IN do que em PB, é natural que também produzam valores menores de H1*-H2*. Portanto, quanto antes o participante tenha adquirido o IN, maior foi a ocorrência de voz crepitante na produção da voz.

A respeito das medidas de f_0 , quanto mais cedo tenha ocorrido o início da aquisição do IN pelo participante, mais baixos foram os valores de f_0 .

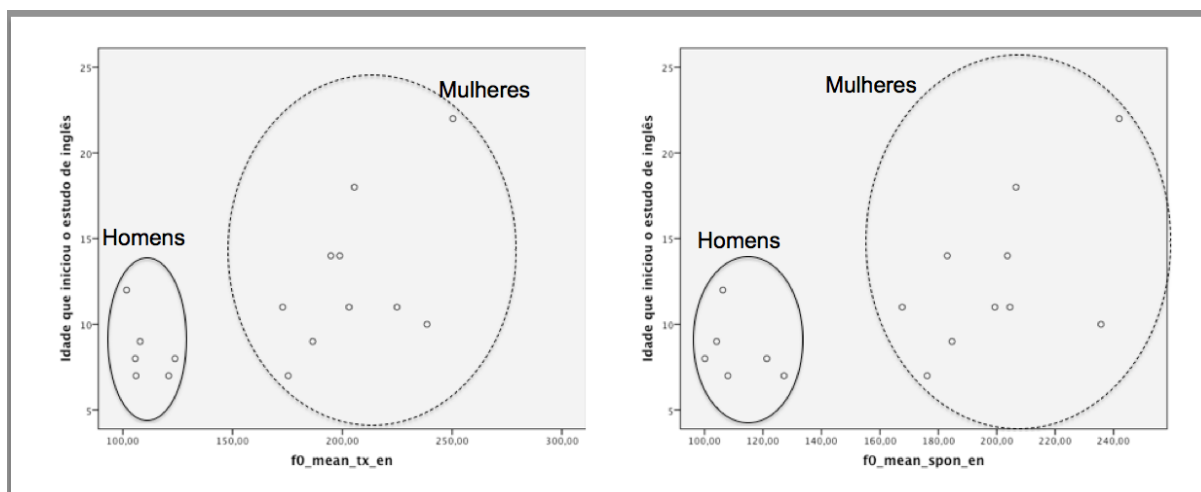


FIGURA 16: GRÁFICO DE CORRELAÇÃO ENTRE MÉDIAS DE F0 EM IN E IDADE DE AQUISIÇÃO

É interessante observar a correlação entre idade de aquisição e média de f0, tanto na leitura de texto, quanto na fala semiespontânea, como ilustrado na Figura 16. Os gráficos de dispersão mostram que, especialmente para as mulheres, quanto mais cedo tenha ocorrido a aquisição do IN, mais baixa é a média de f0.

Para testar a diferença na produção da voz em IN entre os grupos divididos por grau de sotaque estrangeiro, gerados a partir dos julgamentos feitos pelos falantes nativos, foram utilizados os resultados da análise aglomerativa de k-médias descritos na seção 5.2.3.1. Segundo essa análise, os falantes foram divididos em 3 grupos: grupo 1 – pouco sotaque (n=4), grupo 2 – médio sotaque (n=5), e grupo 3 – muito sotaque (n=7). O teste Kruskal Wallis foi rodado para comparar as medidas espectrais de curto e longo termo, bem como as de f0 entre os três grupos nas emissões em IN. Não houve diferença estatisticamente significativa entre os grupos, indicando que a produção da voz em IN não teve relação com o grau de sotaque estrangeiro atribuído pelos falantes nativos.

6.2 EXPERIMENTO DE PERCEPÇÃO

A partir do experimento de percepção, foram geradas médias dos julgamentos igual/diferente feitos pelos ouvintes com relação às vozes dos falantes em PB e em IN. Como reportado no capítulo da Metodologia, os valores gerados

pelo programa TP vão de 0 (vozes iguais em PB e em IN) a 100 (vozes muito diferentes entre o PB e o IN).

De maneira geral, houve bastante variabilidade nos julgamentos feitos pelos ouvintes. Essa variabilidade pode ser verificada a partir do desvio-padrão das médias obtidas pelos julgamentos (TABELA 19):

TABELA 19: MÉDIA E DP DOS JULGAMENTOS DE DIFERENÇAS ENTRE VOZES EM PB E EM IN PELOS OUVINTES

	Média	DP
Fal1	58,0	26,7
Fal2	13,4	27,2
Fal3	27,7	25,9
Fal4	19,5	27,4
Fal5	26,7	31,3
Fal6	15,4	24,5
Fal7	76,0	20,8
Fal8	13,2	24,4
Fal9	57,0	28,6
Fal10	23,6	28,9
Fal11	58,8	26,5
Fal12	17,1	26,5
Fal13	19,8	27,3
Fal14	14,1	27,2
Fal15	39,1	31,1
Fal16	28,8	30,2

A partir dessas médias apresentadas na Tabela 19, verificou-se que os falantes que demonstraram maiores diferenças entre as vozes do PB e do IN foram os falantes 1, 7, 9, e 11 (médias entre 57 e 76). Os falantes 3, 5, 10, 15 e 16 apresentaram diferenças um pouco menores (entre 23,6 e 39,1), enquanto que os falantes 2, 4, 6, 8, 12, 13 e 14 tiveram as menores médias (entre 13,2 e 19,8). Os valores do DP foram em geral bastante altos, sendo os maiores valores atribuídos aos falantes 5, 15 e 16 (acima de 30) e o mais baixo, ao falante 7, que também apresentou a maior diferença entre as vozes do PB e do IN. A variabilidade revelada pelos valores de DP reitera o fato de que, segundo Kreiman e Gerratt (1998), ouvintes individuais são razoavelmente consistentes no julgamento de determinadas qualidades de voz, mas a variância dos julgamentos entre os ouvintes pode ser altíssima, devido a outros fatores como estratégia perceptual e atenção a diferentes aspectos do estímulo.

Para verificar a consistência entre os ouvintes no julgamento das diferenças nas vozes em PB e em IN, foi rodado o teste Pearson com o intuito de conferir a correlação entre os julgamentos. Os ouvintes que obtiveram um maior número de correlações (de oito a onze) uns com os outros foram os de número 1, 2, 6, 7, 8, 10, 11, 12, 13 e 14 ($p < 0,05$). Os outros ouvintes tiveram poucas correlações uns com os outros – de zero a sete. Dessa forma, foram calculadas as médias e DP dos julgamentos feitos por esses dez ouvintes que apresentaram maior número de correlações (TABELA 20):

TABELA 20: MÉDIA E DP DOS JULGAMENTOS DE DIFERENÇAS ENTRE VOZES EM PB E EM IN PELOS OUVINTES QUE MAIS OBTIVERAM CORRELAÇÕES

	Média	DP
Fal1	54,3	27,7
Fal2	8,3	16,6
Fal3	23,4	21,0
Fal4	9,9	19,3
Fal5	32,8	34,1
Fal6	10,0	20,5
Fal7	76,9	17,3
Fal8	5,1	14,6
Fal9	56,9	24,9
Fal10	19,2	27,7
Fal11	62,3	23,8
Fal12	7,1	12,8
Fal13	12,0	21,6
Fal14	5,6	11,6
Fal15	42,3	29,7
Fal16	18,4	20,4

A maioria das médias e DP apresentadas na Tabela 20 foram menores para esse subgrupo de ouvintes. Apenas as médias dos falantes 5, 7, 11 e 15 foram mais altas. Os valores de DP foram também mais baixos a não ser pelos falantes 1 e 5. Assim, para as análises seguintes, serão considerados apenas esses dez ouvintes que apresentaram maior número de correlações entre si.

A fim de agrupar os falantes conforme as médias de diferença entre as vozes em PB e em IN atribuídas pelos ouvintes, foi rodada a análise algomerativa de k-médias, conforme a Tabela 21:

TABELA 21: RESULTADOS DA ANÁLISE AGLOMERATIVA DE K-MÉDIAS PARA O AGRUPAMENTO DOS FALANTES

Número do Falante	Grupo (Cluster)	Distância
1	1	91,83
2	2	32,23
3	2	80,50
4	2	61,28
5	2	128,47
6	2	61,29
7	1	0
8	2	44,37
9	1	98,23
10	2	73,30
11	1	80,52
12	2	40,30
13	2	54,37
14	2	0
15	1	136,58
16	2	77,34

Os resultados apresentados na Tabela 21 mostram que os falantes foram separados em dois grupos: no grupo 1, estão os falantes 1, 7, 9, 11 e 15, e no grupo 2, os falantes 2, 3, 4, 5, 6, 8, 10, 12, 13, 14, e 16. O grupo 1 apresentou as maiores diferenças entre as vozes em PB e em IN, enquanto que o grupo 2 apresentou as diferenças mais sutis entre essas duas línguas. A Tabela 22 traz os resultados do teste Anova gerado pela análise aglomerativa:

TABELA 22: RESULTADOS DO TESTE ANOVA SOBRE A ANÁLISE AGLOMERATIVA DE K-MÉDIAS

	ANOVA					
	Cluster	Erro				
	Quadrado		Quadrado			
	Médio	Df	Médio	df	Z	Sig.
Julg_Ouv1	8665,78	1	488,48	14	17,74	,001
Julg_Ouv2	5557,65	1	193,91	14	28,66	,000
Julg_Ouv6	10857,14	1	351,99	14	30,84	,000
Julg_Ouv7	10902,84	1	566,80	14	19,23	,001
Julg_Ouv8	4921,09	1	260,70	14	18,87	,001
Julg_Ouv10	8744,40	1	459,99	14	19,01	,001
Julg_Ouv11	5871,05	1	287,42	14	20,42	,000
Julg_Ouv12	5863,309	1	372,486	14	15,741	,001
Julg_Ouv13	6169,205	1	188,289	14	32,765	,000
Julg_Ouv14	3371,598	1	163,474	14	20,625	,000

Como mostra a Tabela 22, os valores do quadrado médio do *cluster* revelam que os ouvintes que apresentaram maior poder de discriminação entre os grupos foram os de número 1, 6, 7 e 10. Por outro lado, os ouvintes 2, 13 e 14 tiveram maior poder explicativo para a constituição dos dois grupos de falantes (menores valores do quadrado médio do erro). Os resultados de Z, que é dado pelo quociente entre o quadrado médio do cluster e o quadrado médio do erro, indicam que quem mais contribuiu para a definição dos dois grupos foram os ouvintes 2, 6 e 13.

6.3 RELAÇÃO ENTRE PRODUÇÃO E PERCEPÇÃO

Para tentar estabelecer possíveis relações entre os resultados do experimento de produção e o de percepção, os falantes brasileiros bilíngues foram divididos em dois grupos, conforme os resultados da análise de agrupamento de k-médias (grupo 1: n=5; grupo 2: n= 11). Foi rodado o teste Mann Whitney para verificar diferenças entre os dois grupos de falantes em relação às medidas espectrais de curto e de longo termo, bem como as de f0 (TABELA 23):

TABELA 23: RESULTADOS DO TESTE MANN WHITNEY PARA DIFERENÇAS ENTRE OS DOIS GRUPOS DE FALANTES

Medidas espectrais	H4*-2kHz IN	PPC-Son-IN	D0225kHz-TX IN	D0225kHz-TX PB	D0225kHz-SE PB
Z	-2,209	-1,983	-2,209	-2,549	-1,983
Significância (p)	,027	,047	,027	,011	,047

De acordo com os resultados mostrados na Tabela 23, apenas uma das medidas que apresentou diferença estatisticamente significativa foi da fala semiespontânea, da qual os estímulos do teste de percepção foram retirados. As outras quatro medidas foram obtidas das leituras de frase-veículo e texto. Possivelmente, os falantes produzem características vocais semelhantes tanto na fala semiespontânea quanto nas tarefas de leitura. Contudo, pelo fato de que a fala semiespontânea apresenta maior variabilidade que as tarefas de leitura, as

diferenças entre as línguas tendem a ser minimizadas; por isso a inexistência de diferenças estatisticamente significativas naquele tipo de fala. Outra explicação para a ausência de diferenças entre os grupos de falantes poderia ser o número reduzido de participantes, tanto de falantes quanto de ouvintes.

Para tentar entender os motivos que levaram os ouvintes a perceberem diferenças entre as vozes em PB e em IN dos falantes bilíngues, foram calculadas as médias e DP das medidas de longo termo e de f0 realizadas na fala semiespontânea (TABELA 24):

TABELA 24: MÉDIA E DP DAS MEDIDAS DE LONGO TERMO E DE F0 DA FALA SEMIESPONTÂNEA

Fala SE	Grupo 1 – IN Média (DP)	Grupo 2 – IN Média (DP)	Grupo 1 – PB Média (DP)	Grupo 2 – PB Média (DP)
Razão Alfa	9,19 (3,13)	8,04 (2,53)	9,55 (1,73)	8,23 (2,21)
D0225kHz	17,20 (3,37)	15,38 (2,67)	18,49 (2,54)*	15,72 (2,71)*
D2558kHz	19,80 (3,70)	17,40 (4,44)	18,45 (4,00)	16,23 (3,67)
Média f0	169,95 (54,45)	165,48 (48,67)	174,14 (56,06)	169,33 (48,68)
Extensão f0	307,88 (68,66)	321,84 (71,91)	325,89 (88,49)	308,03 (69,73)

De acordo com a Tabela 24, o grupo 1 é composto pelos falantes que apresentaram mais diferenças entre as vozes em PB e em IN, segundo os ouvintes. O grupo 2, por sua vez, apresentou diferenças mais sutis entre as línguas. Em ambos os grupos, o declínio espectral representado pela razão alfa foi um pouco maior em PB do que em IN. Ao comparar os dois grupos, o grupo 1 apresentou maiores valores na razão alfa que o grupo 2. Isso indica que, de maneira geral, os falantes do grupo 1 produziram mais sopro na voz que os falantes do grupo 2. A mesma tendência ocorreu com a medida D0225kHz, com valores mais altos em PB do que em IN, e mais altos para o grupo 1. Por outro lado, as médias da medida D2558kHz foram mais baixas em PB que em IN, e mais altas para o grupo 1 do que para o grupo 2. Os falantes do grupo 1 produziram mais crepitação na voz que os falantes do grupo 2. Dessa forma, as características vocais de sopro e crepitação parecem ser relevantes para os ouvintes, que mesmo leigos, percebem as diferenças entre as línguas.

Com relação às médias de f0, a tendência foi de *pitch* mais baixo em IN para os dois grupos. Na comparação entre grupos, o grupo 2 produziu *pitch* mais baixo

que os falantes do grupo 1. A respeito da extensão de f0, os grupos se comportaram de maneira diferente nas duas línguas: o grupo 2 obteve maior média que o grupo 1 em IN e o grupo 1 obteve maior média em PB.

Tendo na comparação das medidas espectrais de longo termo quanto nas de f0, os homens e as mulheres estavam misturados nos grupos. Além disso, a proporção de homens e mulheres não era a mesma entre os grupos – grupo 1 com 1,5 mulheres para cada homem e grupo 2, com 1,75 mulheres para cada homem. Por isso, optou-se por separar os homens das mulheres nos grupos 1 e 2:

TABELA 25: MÉDIA E DP DAS MEDIDAS DE ELT E F0 NA FALA SEMIESPONTÂNEA PARA OS HOMENS DOS GRUPOS 1 E 2

Fala SE Homens	Grupo 1 – IN Média (DP)	Grupo 2 – IN Média (DP)	Grupo 1 – PB Média (DP)	Grupo 2 – PB Média (DP)
Razão Alfa	10,52 (3,77)	9,36 (1,24)	10,12 (1,06)	9,32 (1,75)
D0225kHz	19,59 (0,83)	16,63 (1,66)	20,12 (1,90)	16,53 (2,77)
D2558kHz	19,28 (0,08)	20,05 (3,77)	18,60 (2,94)	18,61 (3,11)
Média f0	115,66 (16,32)	108,89 (8,92)	115,26 (9,87)	111,72 (11,13)
Extensão f0	233,12 (8,09)	238,22 (9,97)	229,87 (24,14)	230,95 (22,91)

TABELA 26: MÉDIA E DP DAS MEDIDAS DE ELT E F0 NA FALA SEMIESPONTÂNEA PARA AS MULHERES DOS GRUPOS 1 E 2

Fala SE Mulheres	Grupo 1 – IN Média (DP)	Grupo 2 – IN Média (DP)	Grupo 1 – PB Média (DP)	Grupo 2 – PB Média (DP)
Razão Alfa	8,30 (3,09)	7,29 (2,85)	9,16 (2,22)	7,60 (2,31)
D0225kHz	15,61 (3,59)	14,66 (2,98)	17,40 (2,59)	15,26 (2,78)
D2558kHz	20,15 (5,19)	15,88 (4,28)	18,35 (5,27)	14,86 (3,41)
Média f0	206,15 (29,74)	197,82 (23,54)	213,40 (21,42)	202,25 (20,29)
Extensão f0	357,71 (9,13)	369,62 (35,28)	389,90 (2,45)	352,08 (41,21)

Segundo a Tabela 25⁷⁸, os homens do grupo 1 produziram maior soproidade que os do grupo 2 em ambas as línguas, como se pode perceber pelos valores mais altos da razão alfa e da D0225kHz. O grupo 2, conseqüentemente, produziu mais crepitação que o grupo 1, principalmente em IN, como se nota pelas médias de D2558kHz. Com relação à média de f0, o grupo 1

⁷⁸ Foram destacados nas Tabelas 25 e 26 os valores mais altos de cada medida. São discutidas no texto as comparações entre os grupos e entre as línguas.

produziu *pitch* mais alto que o grupo 2 tanto em IN como em PB. A respeito da extensão de f_0 , o grupo 2 produziu valores mais altos que o grupo 1.

Para as mulheres, segundo a Tabela 26, as participantes do grupo 1 produziram mais soproidade que as do grupo 2. Na comparação entre as línguas, segundo as medidas razão alfa e D0225kHz, houve mais soproidade em PB que em IN. Com relação à medida D2558kHz, os dois grupos produziram maior crepitação em IN que em PB; e o grupo 1, mais crepitação que o grupo 2. A respeito das medidas de f_0 , o grupo 1 apresentou maior média e extensão que o grupo 2; na comparação entre as línguas, os valores foram mais altos em PB que em IN.

Além de os ouvintes quantificarem as diferenças entre as vozes em PB e em IN, eles tinham também que descrevê-las, sempre caracterizando as vozes em IN em relação ao PB. Assim, os dez ouvintes descreveram as vozes em IN dos cinco falantes do grupo 1, os que apresentaram maior grau de diferenças. Para fazer a comparação entre as impressões dos ouvintes e as medidas acústicas dos falantes, foram resumidas todas as descrições feitas e as médias das medidas realizadas nos excertos utilizados no experimento de percepção (QUADRO 6):

Julgamento ouvintes	Medidas Acústicas
Fal1(Homem): mais grave (9), mais baixa (9), mais lento (6), mais inseguro, contido (4), menos variação de tom (4), mais fraca (2).	<u>Média de f_0 mais baixa em IN (105,12 Hz versus 108,28 Hz em PB);</u> extensão de f_0 mais alta em IN (227,40 Hz versus 212,80 Hz em PB); <u>razão alfa mais alta em IN (13,19 dB versus 10,87 dB em PB);</u> <u>D0225kHz mais alta em IN (20,18 dB versus 18,77 dB em PB);</u> D2558kHz mais baixa em IN (19,22 dB versus 20,68 dB em PB); <u>L_{eq} mais alta em IN (44,786 dB versus 43,17 dB em PB).</u>
Fal7(Mulher): mais baixa (14), mais suave (6), mais sussurrada (5), mais aguda (2), mais contida, menos confiante (2).	<u>Média de f_0 mais alta em IN (206,58 Hz versus 205,37 Hz em PB);</u> <u>extensão de f_0 mais baixa em IN (368,15 Hz versus 387,35 Hz em PB);</u> <u>razão alfa mais alta em IN (11,76 dB versus 11,51 em PB);</u> <u>D0225kHz mais alta em IN (19,16 dB versus 19,06 dB em PB);</u> D2558kHz mais alta em IN (14,30 dB versus 12,90 dB em PB); <u>L_{eq} mais baixa em IN (44,080 dB versus 46,92 dB em PB).</u>
Fal 9(Homem): mais alta (18), mais assertiva (4), mais aguda (3), mais confiante (2), mais forte (2), mais expressiva (1).	<u>Média de f_0 mais alta em IN (127,20 Hz versus 122,24 Hz em PB);</u> extensão de f_0 mais baixa em IN (238,85 Hz versus 246,95 Hz em PB); <u>razão alfa mais baixa em IN (7,85 dB versus 9,37 dB em PB);</u> <u>D0225kHz mais baixa em IN (19,01 dB versus 21,46 dB em PB);</u> D2558kHz mais alta em IN (19,34 dB versus 16,52 dB em IN); <u>L_{eq} mais alta em IN (47,33 dB versus 42,21 dB em PB).</u>

Julgamento ouvintes	Medidas Acústicas
Fal11(Mulher): mais alta (18), mais aguda (8), mais confiante (4), maior variação de tom (3), mais rápida (2) mais empolgada (1), mais anasalada (1).	Média de f0 mais baixa em IN (235,68 Hz versus 237,67 Hz em PB); extensão de f0 mais baixa em IN (353,85 Hz versus 390,10 Hz em PB); <u>razão alfa mais baixa em IN</u> (7,35 dB versus 8,88 dB em PB); <u>D0225kHz mais baixa em IN</u> (15,69 dB versus 18,73 dB em PB); D2558kHz mais alta em IN (21,95 dB versus 18,47 dB em PB); <u>L_{eq} mais alta em IN</u> (45,80 dB versus 44,08 dB em PB).
Fal 15(Mulher): mais grave (9), mais baixa (9), menos expressiva (3), mais crepitante (2), menos confiante (2), mais rouca (2), menos variação de tom.	<u>Média de f0 mais baixa em IN</u> (176,19 Hz versus 197,15 Hz em PB); <u>extensão de f0 mais baixa em IN</u> (351,15 Hz versus 392,25 hZ em PB); razão alfa mais baixa em IN (5,79 dB versus 7,10 dB em PB); D0225kHz mais baixa em IN (11,98 dB versus 14,41 dB em PB); <u>D2558kHz mais alta em IN</u> (24, 21 dB versus 23,42 dB em PB); <u>L_{eq} mais alta em IN</u> (45,08 dB versus 46,60 dB em PB).

QUADRO 6: DESCRIÇÕES DAS VOZES EM PB E IN DOS FALANTES DO GRUPO 1⁷⁹

Como os ouvintes descreveram as diferenças de intensidade entre as vozes dos bilíngues – vozes mais altas ou baixas – optou-se por incluir o L_{eq} (*Equivalent Sound Level*) no Quadro X. O L_{eq} é uma média de intensidade do sinal acústico em um determinado tempo. As comparações de intensidade intra-falantes são confiáveis pois os parâmetros de gravação foram exatamente os mesmos nas duas línguas. As medidas acústicas que estão sublinhadas são as que refletem as impressões descritas pelos ouvintes.

O falante 1 obteve média de f0 mais baixa em IN, o que condiz com os nove julgamentos de voz mais grave que recebeu nessa língua. Sua voz também foi julgada como fraca e baixa em IN, características refletidas na razão alfa e na D0225kHz, com maior declínio espectral nessa língua. Essas duas medidas espectrais são correlatas de intensidade; em PB, o declínio espectral foi menor pelo fato de haver mais energia concentrada na faixa mais alta de frequência, caracterizando uma voz mais alta e/ou forte na L1 do falante.

A falante 7 teve a voz em IN descrita como mais aguda que no PB, o que foi constatado pelas médias de f0, com a média da produção em IN mais alta que a média em PB. A extensão de f0 mais baixa em IN foi condizente com a característica de fala mais contida em IN. Os atributos “mais suave” e “mais sussurrada” foram confirmados pela razão alfa e pela D0225kHz, ambas com

⁷⁹ Ao lado dos julgamentos dos ouvintes na primeira coluna, entre parênteses, está o número de vezes que aquela descrição foi utilizada para caraterizar as vozes dos falantes.

valores mais altos em IN, indicando voz mais soprosa nessa língua. Por fim, a média mais baixa de L_{eq} também confirmou a atribuição de voz mais baixa e/ou fraca em IN que em PB.

O falante 9 teve sua voz em IN julgada como mais alta e mais aguda, características que foram confirmadas pela maior média de f_0 e pela L_{eq} mais alta em IN que em PB. A razão alfa e a D0225kHz mais baixas em IN foram condizentes com o julgamento de voz mais alta nessa língua.

A voz em IN da falante 11 foi julgada como mais alta que em PB, o que foi constatado pela L_{eq} . Essa característica também foi reiterada pela razão alfa e pela D0225kHz, que foram mais baixas em IN. Contudo, o atributo “mais aguda” não foi confirmado pela média de f_0 , que foi mais alto em PB que em IN.

Por fim, a falante 15, a quem foi atribuída voz mais grave em IN, teve média de f_0 mais baixa nessa língua. Sua voz em IN também foi caracterizada como tendo menos variação de tom, confirmada pela menor extensão de f_0 que em PB. O fato de ela ser julgada como mais baixa em IN não foi constatado pela razão alfa, D0225kHz ou pela L_{eq} . A D2558kHz foi mais alta em IN, indicando voz mais crepitante nessa língua – característica que foi percebida em sua fala.

6.4 DISCUSSÃO

Nesta seção, são discutidos os resultados dos testes de produção e de percepção, bem como a relação entre eles, no que concerne à produção da voz e a percepção de sua qualidade por falantes e ouvintes brasileiros bilíngues.

Primeiramente, a respeito dos resultados do experimento de produção, ficou evidente que os brasileiros bilíngues que participaram deste estudo, de maneira geral, produziram vozes em IN com algumas características distintas das vozes produzidas em PB. É bastante importante ressaltar que quando se reporta diferenças entre as línguas, estamos utilizando uma como referência da outra. Portanto, quando apontamos uma tendência de voz mais soprosa em uma língua, é porque a outra apresenta menos características acústicas de voz soprosa;

utilizamos o *continuum* de fonação de Ladefoged (1971), pois consideramos que a soproidade e a crepitação são atributos gradientes, e não discretos.

Na tarefa de leitura do texto, os brasileiros produziram vozes mais soprosas em PB e vozes mais crepitantes em IN, assim como indicaram as três medidas espectrais de longo termo. Já na fala semiespontânea, houve diferença apenas na medida de declínio espectral que indica a crepitação na voz – a D2558kHz. Os falantes produziram vozes mais crepitantes em IN do que em PB. Esses resultados estão em consonância com os de Ng *et al.* (2012), que analisaram a qualidade de voz em Cantonês e IN por chineses bilíngues proficientes no que concerne à tarefa de leitura de texto. Seus resultados mostraram que os chineses produziram vozes mais soprosas em cantonês que em IN. Da mesma forma, Harmegnies e Landercy (1985), comparando bilíngues de francês e holandês, Harmegnies *et al.* (1989) e Bruyninckx *et al.* (1994) comparando bilíngues de catalão e espanhol, reportaram diferenças na qualidade de voz por medidas de espectro de longo termo entre as línguas. Contudo, esses estudos não contaram com a análise de amostras de fala semiespontânea, utilizando somente a tarefa de leitura de texto em seus experimentos. Como aponta Byrne *et al.* (1994), medidas espectrais de longo termo são certamente influenciadas pelo tipo de material de fala analisada, mas esses autores discutem apenas as diferenças entre listas de palavras ou conjuntos de frases e pequenos textos. Os nossos resultados, que indicam diferenças estatisticamente significativas em todas as medidas espectrais de longo termo na tarefa de leitura, parecem indicar que os bilíngues produzem vozes distintas entre as línguas, como se houvesse um “*reading mode*” específico para cada uma.

Essa tendência de produzir vozes menos soprosas, e muitas vezes mais crepitantes em IN tanto para os homens, mas principalmente para as mulheres, vai de encontro com pesquisas recentes relacionadas aos tipos de fonação e sexo. Podesva (2013) é um exemplo de estudo que mostra que as mulheres estadunidenses têm produzido altos índices de voz crepitante em sua fala. Esse autor cita uma série de outras pesquisas que têm indicado a mesma tendência, inclusive apontando as possíveis motivações para tal fenômeno. Dentre as impressões que uma voz mais crepitante pode causar estão: assumir postura mais autoritária, mas sem soar agressiva; estar em posição social ascendente, ser

profissional, urbana; ser capaz de competir à altura com os homens. Segundo Podesva (2013), “*de acordo com essa interpretação, portanto, o significado social da voz crepitante está enraizada na sua semelhança com a voz dos homens*”⁸⁰. Ele ainda sugere que as altas taxas de voz crepitantes encontradas na fala feminina podem ser consequência do abaixamento do *pitch*.

Na comparação entre as tarefas, houve diferença estatisticamente significativa em algumas medidas espectrais em cada língua. A tarefa de leitura de texto apresentou diferença apenas na razão alfa no IN em comparação à fala semiespontânea, com a presença de maior sopro na leitura de texto. No PB, contudo, as três medidas apresentaram diferenças entre as tarefas: mais sopro na leitura de texto e mais crepitação na fala semiespontânea. O fato de haver mais diferenças nas vozes entre as tarefas em PB que em IN converge com os resultados de Bruyninckx *et al.* (1994), que mostraram que há uma maior coerência ou regularidade na língua não dominante e uma maior variabilidade na língua dominante.

Além da dominância linguística e do tipo de tarefa provocarem diferenças nas vozes dos bilíngues entre as emissões em PB e em IN, podem também estar envolvidas no fenômeno a questão da personalidade e da emoção. Sabe-se que a voz é altamente condicionada tanto pelos estados emocionais transientes quanto pelas características identitárias mais estáveis (KREIMAN e SIDTIS, 2011). Os atributos vocais do falante podem revelar características físicas (altura, beleza, idade), psicológicas (inteligência, extroversão, dominância), emocionais, bem como relacionadas à masculinidade ou feminilidade. Some-se a isso, a qualidade de voz também tem motivação cultural, podendo variar entre as línguas por conta de estereótipos vocais existentes em diferentes países. Todas essas explicações podem dar conta das diferenças produzidas pelos bilíngues da presente pesquisa.

As características de personalidade e emoção, que explicam as diferenças na qualidade de voz dos falantes bilíngues entre o PB e o IN, podem também dar conta da distinção entre as medidas de f_0 . Houve diferenças entre as médias de f_0 apenas na leitura do texto, sendo essas mais baixas em IN que em PB, ao

⁸⁰ According to this interpretation, then, the social meaning of creaky voice is rooted in its resemblance to men's voices.

contrário das de extensão, que foram mais altas em IN. Esses resultados estão de acordo com os de Todaka (1995) e Altenberg e Ferrand (2006) com relação às médias produzidas em L1 e L2 de seus participantes. Porém, os resultados de Ng *et al.* (2012) mostraram o oposto: os bilíngues cantoneses produziram maiores médias de f0 em IN que na sua L1. Pépiot (2013), que não comparou bilíngues, mas franceses falando francês a estadunidenses falando IN, constatou que os franceses obtiveram médias de f0 mais altas que os estadunidenses. As diferenças de f0 entre línguas em falantes bilíngues são bastante plausíveis, pois, segundo Keating e Kuo (2010), os aprendizes de línguas geralmente entendem que determinados ajustes de *pitch* são apropriados para uma língua e não para outra.

Quanto às medidas de f0 na fala semiespontânea, não houve diferenças estatisticamente significativas entre as emissões em IN e em PB. Isso aconteceu possivelmente pela natureza da tarefa: a fala semiespontânea, por ser menos controlada, foi produzida pelos falantes com f0 com maior variância. Da mesma forma, na comparação entre as tarefas, não houve diferenças entre a leitura de texto e a fala semiespontânea. Mesmo assim, a tendência de médias mais baixas de f0 em IN foi também constatada na fala semiespontânea.

Segundo Kreiman e Sidtis (2011), com base em pesquisas que demonstram a percepção da personalidade pelo ouvinte por meio de características vocais, os efeitos da f0 variam bastante: um *pitch* mais alto pode tanto estar associado a uma maior extroversão, assertividade, autoconfiança e competência quanto à imaturidade, instabilidade emocional, nervosismo e inveracidade. A variabilidade de *pitch* pode estar associada à percepção de personalidade mais dinâmica, extrovertida, benevolente e submissiva. Assim, uma possível explicação para a produção de f0 mais baixa em IN que em PB pode ser de que os brasileiros bilíngues têm mais autoconfiança e são mais extrovertidos na sua L1.

As diferenças na produção da voz entre homens e mulheres no que concerne às medidas espectrais e de f0 já eram esperadas, tanto por razões fisiológicas quanto culturais. A respeito das medidas espectrais, Klatt e Klatt (1990) afirmam que homens e mulheres variam principalmente em dois parâmetros acústicos que indicam sopro: a relativa amplitude do primeiro harmônico (H1) e a quantidade de ruído de aspiração na região do F3 (entre 2400 a 3000Hz);

esses autores ressaltam que as mulheres produzem vozes bastante mais soprosas que homens.

Na presente pesquisa, houve diferenças estatisticamente significativas nas medidas espectrais de curto termo $H1^*-H2^*$ e 2kHz-5kHz em PB e IN entre homens e mulheres, resultados que vão de encontro às características apontadas por Klatt e Klatt (1990). Os homens apresentaram valores mais baixos de $H1^*-H2^*$ e mais altos de 2kHz-5Kz que as mulheres, tanto em PB quanto em IN. Com relação às medidas espectrais de longo termo, houve diferença estatisticamente significativa apenas na D2558kHz na tarefa de leitura de texto em PB. Da mesma forma, os homens apresentaram valores mais altos de declínio espectral que as mulheres, caracterizando vozes mais crepitantes para eles e mais soprosas para elas.

As diferenças nas medidas de f_0 entre homens e mulheres foram verificadas nas leituras de texto e na fala semiespontânea, tanto no IN quanto no PB. Os estudos de Altenberg e Ferrand (2006), Ng *et al.* (2012) e Pépiot (2014) também reportaram diferenças nas medidas de f_0 entre homens e mulheres que participaram de suas pesquisas, tendo as mulheres médias e extensões de f_0 mais altas que os homens.

A relação entre a produção da voz em IN e as características de aquisição dessa língua estrangeira pelos brasileiros bilíngues foi verificada no que concerne ao tempo de experiência em países que usam o IN como L1 e também pela idade de aquisição desta L2 pelo falante. Houve diferença estatisticamente significativa apenas para a variável idade de aquisição. Ficou evidenciado que quanto mais cedo tenha ocorrido a aquisição do IN, mais baixos foram os valores de $H1^*-H2^*$ e medidas de f_0 em ambas as tarefas de leitura e de fala semiespontânea em IN. Para explicar essa tendência, o trabalho de Podesva (2013) pode ser uma alternativa no que diz respeito, principalmente, à voz das participantes mulheres. Parece que, quanto mais jovem a bilíngue tenha aprendido o IN, mais graves e crepitantes são suas vozes nessa L2 – característica de muitas mulheres estadunidenses jovens. Em relação ao tempo de experiência em países de L1, variável que não mostrou correlação com as características das vozes dos bilíngues brasileiros, como aponta Flege (2009), não somente o tempo como também a qualidade do input ao qual o aprendiz é exposto e a frequência da

interação feita em L2 exercem influência no processo de aquisição. Outro fator que parece não exercer influência na produção da voz em IN dos falantes bilíngues é o grau de sotaque, que foi atribuído pelos falantes nativos dessa L2. Segundo essa análise, o fato de um falante apresentar maior grau de sotaque estrangeiro não necessariamente implica na produção de determinadas características vocais como soproidade, crepitação, *pitch* mais alto ou baixo.

Como explica Moyer (2004), há outros fatores para além da idade de aquisição e tempo de experiência que influenciam a aprendizagem da L2. Ela relata que a intensidade da motivação em termos do alcance de uma boa proficiência ou em termos profissionais são fatores que influenciam fortemente a aprendizagem da L2. Além desses citados, as oportunidades de contato naquela L2, a atitude em relação à cultura do país em questão e a intenção em permanecer nele têm peso na progressão durante o processo de aprendizagem. Pode ser que os participantes do presente estudo tenham sido influenciados de maneira mais contundente por esses fatores sociais e afetivos envolvidos na aquisição da L2, principalmente pela qualidade de voz ser crucial para veicular atributos paralinguísticos na comunicação. Contudo, como não é do escopo de nossa pesquisa, não foram coletados dados de motivação e atitude para com o IN e, por isso, não se pode afirmar inequivocadamente que a qualidade de voz seja realmente afetada por tais fatores sociais e afetivos.

Em se tratando dos resultados do experimento de percepção, houve relativa variabilidade no julgamento das vozes produzidas pelos falantes bilíngues, feito pelos ouvintes leigos também bilíngues. Na verificação da consistência dos julgamentos entre os ouvintes, foi constatado que dez dos 14 ouvintes produziram julgamentos compatíveis entre si a respeito das diferenças de qualidade de voz produzidas pelos falantes entre o PB e o IN. Acredita-se que a tarefa do experimento de percepção tenha contribuído para a confiabilidade dos julgamentos entre esses dez ouvintes, pois consistia em comparar emissões em IN em relação às em PB do mesmo ouvinte. Segundo Kreiman *et al.* (2007), quando é fornecido ao ouvinte um estímulo de referência para o seu julgamento, a dependência de padrões internos de qualidade de voz é eliminada e a concordância entre os ouvintes aumenta. Além disso, o fato de os ouvintes terem familiaridade com

ambas as línguas dos falantes também contribuiu para a confiabilidade dos julgamentos. Assim, a partir dos julgamentos feitos pelos dez ouvintes mais consistentes entre si, os 16 falantes foram separados em dois grupos estabelecidos pela quantidade de diferenças entre as vozes em PB e em IN: o grupo que apresentava maior distinção na voz em IN em comparação à voz em PB, e o grupo que apresentou menor diferença entre as línguas.

Na tentativa de estabelecer possíveis relações entre as características acústicas das vozes produzidas pelos falantes e o julgamento atribuído pelos ouvintes, buscou-se comparar as medidas espectrais e de f_0 entre os dois grupos de falantes (os que tinham sido julgados como tendo vozes mais diferentes – o grupo 1 – e os que tinham vozes mais parecidas entre o PB e o IN – grupo 2). Os resultados mostraram baixa correlação entre os julgamentos e as medidas espectrais e de f_0 na fala semiespontânea, da qual foram selecionadas as amostras para o experimento de percepção. A única medida que apresentou diferença estatisticamente significativa na fala semiespontânea foi a D0225kHz. Os falantes do grupo 1 tiveram uma maior diferença entre as médias do IN e do PB, produzindo vozes mais soprosas em PB que em IN (atestado pelo declínio espectral mais acentuado em PB), enquanto que o grupo 2 apresentou diferença mínima entre as línguas. Outra característica fornecida pela medida D0225kHz tem relação com a crepitação: segundo Keating e Garellek (2015), a voz crepitante também apresenta harmônicos mais fortes nas regiões mais altas do espectro (até os 5kHz). Portanto, o grupo 1 parece ter produzido uma maior diferença entre o PB e o IN nas características de soprosidade e de crepitação. Esses atributos certamente foram decisivos e percebidos pelos ouvintes como pistas durante os julgamentos no experimento de percepção.

Na descrição das diferenças detectadas pelos ouvintes para os falantes do grupo 1, foi verificada uma relativa consistência entre os ouvintes, mas nem sempre houve correspondência entre essas descrições e as medidas realizadas nas vozes de cada falante. Para quatro dos cinco falantes, as descrições de *pitch* realizadas pelos ouvintes corresponderam às diferenças médias de f_0 entre o PB e o IN. Da mesma forma, a razão alfa e a D0225kHz também alcançou alta taxa de correspondência entre o que foi percebido pelos ouvintes e o que foi realizado

pelos falantes. Em três dos cinco falantes, as diferenças de intensidade da voz foram adequadamente captadas pelos ouvintes. O mesmo não ocorreu para extensão de f_0 , pois apenas dois dos cinco falantes apresentaram correspondência entre os julgamentos e as médias realizadas. Por fim, em um caso apenas, a D2558kHz correspondeu ao julgamento feito pelos ouvintes. Três dos cinco falantes produziram vozes mais agudas em IN que em PB, contrariando a tendência de vozes com *pitch* mais baixo em IN que ocorreu tanto na tarefa de leitura (com diferenças estatisticamente significativas) como na fala semiespontânea. Os atributos de sopro e crepitação, assim como os de intensidade, não foram os mesmos para todos os falantes; cada um deles parece produzir a voz em LI utilizando características vocais distintas. Como os ouvintes colocaram, alguns falantes parecem mais tímidos em IN, enquanto outros parecem mais confiantes e falam mais alto nessa língua, em comparação com o PB. Uns falantes são mais expressivos, outros produzem vozes mais monótonas em IN que em PB. Essas diferenças mostram que, possivelmente, os falantes fazem uso de diferentes características vocais por diferentes motivações.

Em suma, os brasileiros bilíngues falantes de PB e IN parecem produzir vozes com algumas diferenças entre as línguas. De maneira geral, em termos fonatórios, os falantes produzem vozes mais suaves e soprosas em PB em comparação ao IN, enquanto que nessa L2, produzem vozes um pouco mais graves e com maior índice de crepitação. O tom, em IN, tende a ser mais baixo que no PB. Essas diferenças, no entanto, são mais evidentes na tarefa de leitura de texto do que na fala semiespontânea, pois aquela demanda uma produção mais controlada da fala. Sugere-se portanto que, por conta da maior diferença entre as línguas, os falantes tivessem um “*reading mode*” para cada língua, e que assim, ajustavam suas vozes a depender de qual língua estivessem falando. Por outro lado, na fala semiespontânea, o bilíngue pode ter que concentrar-se mais no planejamento do conteúdo semântico e sintático do enunciado a ponto de deixar de lado a preocupação em soar diferente em IN. Outro ponto bastante interessante do presente estudo foi a correlação (ou falta dela) entre as características vocais e as variáveis relacionadas ao processo de aquisição do IN e ao grau de sotaque estrangeiro atribuído pelos falantes nativos dessa L2. A única variável que mostrou

correlação entre características fonatórias e de f_0 foi a idade de aquisição do IN. Na literatura a respeito da aquisição da L2, essa não é uma questão resolvida, porém, é recorrente a ideia de que quanto mais cedo se aprenda a língua em questão, melhor é o desempenho e o sotaque em termos do padrão nativo. Os falantes desta pesquisa que aprenderam IN mais cedo, produziram vozes mais graves e mais crepitantes que os que aprenderam mais tarde. Contudo, O fato de o falante ter morado por anos em países onde se fala IN ou de ele produzir essa L2 com forte sotaque estrangeiro não tiveram correlação alguma com as diferenças entre as línguas. Ficou também evidenciado que, no que concerne à percepção da mudança de qualidade de voz pelos falantes bilíngues, os ouvintes leigos percebem as diferenças e as quantificam de maneira relativamente uniforme. De 14 ouvintes, dez fizeram julgamentos semelhantes das vozes em PB e em IN dos falantes e, em muitos casos, descreveram as diferenças de maneira análoga também. Certamente não há unanimidade nos julgamentos, mas as propriedades fonatórias e de f_0 mais discrepantes entre as línguas foram percebidos pelos ouvintes.

7. CONCLUSÃO

Esta tese teve como objetivo principal verificar possíveis mudanças na produção da voz de bilíngues brasileiros quando falam IN em comparação a quando falam PB, bem como a percepção das diferenças de qualidade voz por ouvintes leigos. A partir dos resultados apresentados e da discussão em torno deles, ficou evidenciado que os participantes desta pesquisa produzem algumas diferenças entre as vozes no PB e no IN e que, de certa forma, essas diferenças são percebidas pelos ouvintes leigos também bilíngues. Dentre as características que mais se distinguem entre as línguas estão as medidas espectrais de longo termo, especialmente a razão alfa e a D0255kHz, e a média de f0. As diferenças detectadas entre o PB e o IN pelas medidas acústicas variaram principalmente em função da tarefa desempenhada, do sexo, e da idade de aquisição do IN.

Em resumo, as perguntas de pesquisa, hipóteses e resultados são apresentados no Quadro 7:

Pergunta de pesquisa	Hipótese	Resultado
1. Há diferenças na fonação (medidas de curto termo) entre PN e IN?	H1. Não há diferenças na fonação (medidas de curto termo) entre o PB e o IN.	Não houve diferenças entre PB e IN. Portanto, H_0 mantida.
2. Há diferenças nas características espectrais de longo termo entre PB e IN? Há diferenças entre as tarefas?	H2.1 Há diferenças nas medidas espectrais entre as línguas. H2.2 Há diferenças entre as medidas espectrais entre as tarefas.	1. Houve diferenças na razão alfa, D0225kHz e D02558kHz entre as línguas na tarefa de leitura e na D2558kHz na fala semiespontânea. 2. Houve diferenças na razão alfa, D0225kHz e D02558kHz entre tarefas no PB, mas apenas na razão alfa no IN.
3. Há diferenças nas medidas de f0 entre o PB e o IN? Há diferenças entre as tarefas?	H3.1 Há diferenças nas medidas de f0 entre PB e IN. H3.2 Há diferenças de f0 entre as tarefas.	1. Houve diferenças nas medidas de f0 entre o PB e o IN apenas na tarefa de leitura de texto. 2. Não houve diferenças nas medidas de f0 entre as tarefas em PB ou IN.
4. Há diferenças nas medidas espectrais de curto e longo termo e de f0 entre homens e mulheres?	H4. Há diferenças nas medidas espectrais e nas medidas de f0.	Houve diferenças nas medidas de curto termo H1*-H2* e 2k-5kHz em PB e em IN, e na de longo termo D2558kHz na leitura de texto em PB. Houve diferenças nas medidas de f0 em PB e em IN na leitura de texto e fala semiespontânea.

Pergunta de pesquisa	Hipótese	Resultado
5. As medidas espectrais de curto e longo termo e de f0 estão associadas à idade de aquisição do IN e experiência em países de LI?	H5. Há correlação entre as medidas espectrais de curto e longo termo e fatores como idade de aquisição do IN e permanência em países de LI.	Houve correlação entre H1*-H2* e as medidas de f0 e a idade de aquisição do IN.
6. Os ouvintes leigos percebem diferenças entre as vozes em PB e em IN de brasileiros bilíngues?	---	A maior parte dos ouvintes (10/14) atribuiu julgamentos consistentes entre si em relação às diferenças de qualidade de voz produzidas pelos falantes.
7. Como os ouvintes quantificam e descrevem as diferenças nas vozes entre o PB e o IN dos brasileiros bilíngues?	—	Os ouvintes utilizam descritores acústicos (baixo e alto, grave e agudo), e de outra natureza (suave, fraca), além de descritores psicológicos e de identidade (empolgado, tímido, assertivo, confiante, inseguro).

QUADRO 7: RESUMO DAS PERGUNTAS DE PESQUISA, HIPÓTESES E RESULTADOS

A qualidade de voz é a matéria-prima da prosódia, sendo responsável por veicular tanto características linguísticas como paralinguísticas na comunicação. Suas funções, especialmente no domínio paralinguístico – o mais importante para este trabalho – envolvem a expressão da emoção e a intenção do falante, bem como fornecem características psicológicas e de personalidade dos falantes. Como relata Crystal (1975), é através da paralingua que o falante sinaliza intenções como persuadir e alegrar, além de expressar estados psicológicos como dominância, competência e submissão. Em pesquisas sobre os efeitos da língua na expressão da personalidade em falantes bilíngues, há evidências que, a depender da língua que utilize, o falante muda suas características de personalidade principalmente para adequar-se aos padrões culturais adotados por determinada sociedade. Chen e Bond (201) afirmam que características de personalidade expressas pela fala podem ter origem em diferentes papéis sociais que o falante desempenha entre as línguas. Segundo esses autores, as línguas que os bilíngues falam ativam diferentes comportamentos culturais específicos de cada língua. Outro fator que parece estar associado às mudanças de personalidade em função da língua é a competência linguística, de forma que quanto mais fluente o bilíngue seja, menos diferenças de personalidade ele produza. Os resultados de Chen e Bond (2010),

que investigaram a percepção da personalidade e a acomodação cultural em diálogos entre chineses bilíngues e estadunidenses, mostraram que estes são percebidos como tendo personalidade mais extrovertida, emocionalmente mais estável, mais aberta e mais assertiva que os chineses. Portanto, é possível especular que, para os falantes bilíngues do presente estudo, as mudanças na voz entre o PB e o IN podem ter origem nas diferenças culturais que existem entre a sociedade brasileira e a estadunidense. Nesse raciocínio, a motivação para as mudanças de qualidade de voz dos brasileiros pode ser uma adaptação às características de personalidade atribuídas aos estadunidenses.

Contudo, pelo baixo número de pesquisas sobre a qualidade de voz e a fala bilíngue, não é possível estabelecer interpretações inequívocas sobre as motivações que levam um falante a produzir vozes diferentes em línguas diferentes. Essa dificuldade se dá principalmente pela pouca atenção que aspectos prosódicos têm recebido na pesquisa sobre a aquisição da L2. Pouco se sabe sobre como um aprendiz de L2 adquire elementos prosódicos da L2 - ou mesmo se esses elementos são “ensináveis” e “aprendíveis”. Por isso, ainda não está claro para aprendizes e professores onde reside a dificuldade, se na produção dos elementos prosódicos, na percepção dos mesmos, ou se nas duas esferas. Isso se reflete na ausência de materiais didáticos que proponham o aumento da consciência do aprendiz de L2 sobre a importância das características prosódicas da L2 nos atos de comunicação. Mais especificamente com relação às pesquisas sobre qualidade de voz, o número também reduzido de estudos sobre a voz dos bilíngues ainda conta com a grande variedade de formas de medir ou avaliar o fenômeno, de encaminhamentos metodológicos, e de tratamento e apresentação dos resultados. Ademais, as tarefas desempenhadas pelos participantes das pesquisas variam bastante, dificultando a comparação entre os resultados e possíveis explicações para o fenômeno. As principais diferenças entre o presente estudo e os outros previamente realizados se deram nas medidas acústicas de longo termo, nos tipos de tarefa e no tratamento estatístico dos dados. Contudo, foi possível estabelecer correspondências entre as pesquisas: houve similaridades entre os nossos resultados e os de Todaka (1995), Mendoza *et al.* (1996), Altenberg e Ferrand (2006), Ng *et al.* (2012), e Camargo *et al.* (2013) em termos de

medidas espectrais de longo termo e de f_0 . Camargo *et al.* (2013) foi o único estudo que utilizou tarefas de leitura de texto e fala semiespontânea e, portanto, foi possível comparar as diferenças entre tarefas obtidas em nosso estudo e no estudo deles. Assim como foram mostrados pelos nossos resultados, eles também verificaram diferença estatisticamente significativa apenas nas medidas espectrais de longo termo na fala semiespontânea, enquanto que na tarefa de leitura houve diferenças entre as medidas de f_0 , intensidade e declínio espectral.

Não foi possível dialogar com outros estudos no que concerne ao experimento de percepção realizado nesta tese, pois, até onde se saiba, não há estudos relacionados à qualidade de voz em bilíngues nos mesmos moldes. O *design* deste experimento foi pensado para responder uma das perguntas principais da pesquisa: o ouvinte bilíngue PB/IN leigo percebe mudanças de qualidade de voz em falantes também bilíngues? Assim como evidencia a literatura, os ouvintes extraíram as características acústicas da voz, como *pitch*, intensidade e qualidade de voz, pela comparação entre as emissões em PB e em IN, bem como características emocionais e de personalidade, como empolgação, assertividade e autoconfiança. Os resultados mostraram que, embora com alguma variabilidade, os ouvintes apresentaram julgamentos parecidos entre si. Ademais, as descrições das diferenças feitas pelos ouvintes, em grande parte, tiveram relação com as medidas acústicas realizadas na produção dos bilíngues.

Assim, entende-se que o presente estudo, inédito pelo desenho metodológico e pela relação entre a produção da voz e a percepção de sua qualidade dos falantes bilíngues, traz contribuições tanto para a área da qualidade de voz quanto para os estudos sobre a aquisição de L2. Esta pesquisa mostrou que há correlação entre a voz que falantes bilíngues produzem e a qualidade de voz que ouvintes bilíngues leigos percebem, relação que nenhum outro estudo havia explorado.

Atualmente, com o avanço do desenvolvimento da tecnologia da fala, fica ainda mais nítida a necessidade da realização de pesquisas sobre a produção da voz e a percepção da sua qualidade. No caso de uma tecnologia que permita a tradução subsequente da fala de uma língua para outra, é necessário que sejam levantadas questões acerca dos seguintes pontos: os efeitos da utilização de

determinados modelos computacionais; como medir se alguém soa a mesma pessoa entre as línguas; quais os impactos que as diferenças entre fala natural e a sintetizada pode ter, entre outros. Para que uma pessoa possa obter uma tecnologia personalizada de tradução de voz que a permita soar em uma L2 igualmente como soa na sua L1, primeiramente é preciso que se entenda como se dá a percepção de aspectos inerentes à voz humana e como o modelo computacional daria conta de atributos tão sutis que os falantes veiculam por sua fala.

Com relação às implicações pedagógicas desta pesquisa, parece ser pouco eficaz tentar “ensinar” como seria uma qualidade de voz adequada ao IN como L2. Isso acontece pois a qualidade de voz, pelo menos no domínio paralinguístico, está vinculada a uma série de outros aspectos, como as questões de identidade, emoção e atitude expressadas na fala. Assim, não há uma qualidade de voz que seja “correta” para uma língua, pois embora ela tenha um aspecto social no que diz respeito às características vocais de uma comunidade linguística, ela é altamente idiossincrática e variável. Além do mais, em tempos em que o IN é considerado uma língua franca ou língua internacional, não há mais no processo de ensino/aprendizagem desta L2 uma preocupação exacerbada em soar estadunidense, inglês ou australiano – preza-se principalmente a inteligibilidade. Outro argumento, alinhado ao conceito de IN como língua franca, está a desmitificação do falante nativo, que não mais figura como o modelo de falante e de fala em L2. Contudo, um trabalho de conscientização a respeito dos aspectos da produção da voz e da percepção de sua qualidade pode ser realizado em sala de aula, a fim de discutir as diferentes nuances da voz entre as línguas e como interpretamos determinados atributos vocais em culturas diferentes.

É evidente que se alguns atributos da presente pesquisa tivessem sido diferentes ou se fossem tomadas diferentes decisões metodológicas, poderiam ser esperados resultados distintos. Primeiramente, se o número de participantes do experimento de produção fosse maior, com quantidade equilibrada de homens e mulheres, os testes estatísticos poderiam trazer resultados mais contundentes. A principal dificuldade do presente estudo foi recrutar indivíduos que atendessem os requisitos de participação, principalmente pela coleta de dados ter sido realizada

nos Estados Unidos. Além do número de participantes, o perfil dos indivíduos no que diz respeito à aquisição do IN foi bastante heterogêneo, levando em consideração fatores como idade de aquisição, experiência em países de IN, e grau de proficiência. Especula-se que se a coleta tivesse acontecido no Brasil, o grupo de participantes pudesse ser mais homogêneo nesses termos, trazendo resultados distintos. Outra possibilidade metodológica diz respeito à escolha das medidas acústicas para analisar os dados. Pode ser que outras medidas utilizadas pela literatura captem características distintas da qualidade de voz que não foram contempladas nesta pesquisa. Alguns exemplos são as medidas de amplitude do primeiro pico espectral, que revela características da tensão glotal, e a média de energia espectral, que tem relação com a tensão laríngea. Por fim, se outras amostras de fala tivessem sido usadas como estímulos do experimento de percepção da voz por ouvintes leigos, os julgamentos poderiam ser distintos.

Como sugestão para pesquisas futuras, poderiam ser feitas investigações com a participação de grupos mais homogêneos de bilíngues em termos de características de aquisição de L2, ou com bilíngues de outras línguas. Seria também interessante coletar dados de estadunidenses que falassem PB e comparar suas vozes entre as línguas. Com relação ao experimento de percepção, poderiam ser realizados julgamentos das vozes por pessoas experientes na avaliação vocal, ou por meio de outros modelos teóricos ou métodos distintos. Além disso, poderiam ser julgadores das vozes ouvintes falantes nativos de IN.

Em conclusão, a presente pesquisa vem a contribuir com a área dos estudos relacionados à produção da voz e percepção de sua qualidade no que concerne às diferenças que bilíngues apresentam em sua fala na utilização de uma ou de outra língua. Parece que a impressão de vozes distintas, comumente percebidas por aprendizes de L2 ou bilíngues no dia-a-dia, realmente existe para algumas pessoas. Ficou claro neste estudo que nem todos os bilíngues produzem vozes diferentes em línguas distintas, e quando o fazem, não apresentam necessariamente sempre as mesmas características que os outros bilíngues. As mudanças na voz acontecem por vários motivos e de diferentes formas para os falantes bilíngues.

REFERÊNCIAS

ABERCROMBIE, D. **Elements of general phonetics**. Edinburgh: Edinburgh University Press, 1967.

ALKU, P. Glottal inverse filtering analysis of human voice production – A review of estimation and parameterization methods of the glottal excitation and their applications. **Sadhana** **36** (5), 623–650, 2011.

ALTENBERG, E. P.; FERRAND, C. T. Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. **Journal of Voice**, 20(1), 89–96, 2006.

ANDRUSKI, J.; e RATLIFF, M. Phonation types in production of phonological tone: The case of Green Mong. **Journal of the International Phonetic Association**, **30**, 37–61, 2000.

BAPTISTA, B. O. The Acquisition of English Vowels by BrazilianPortuguese Speakers. **Advanced Research English Series – ARES**, v. 6. Florianópolis: Pós-Graduação em Inglês, Universidade Federal de Santa Catarina, 2000.

BAPTISTA, B. O.; SILVA FILHO, J. L. A. The influence of markedness and syllable contact on the production of English final consonants by EFL learners. In J. Leather e A. James (Eds.), **New Sounds 1997: Proceedings of the Third International Symposium on the Acquisition of Second Language Speech** (pp. 26-34). Klagenfurt: University of Klagenfurt, 1997.

BARBOSA, P.A. Detecting changes in speech expressiveness in participants of a radio program. In: **Proceedings. of Interspeech 2009 - Speech and Intelligence**. Londres: Causal Productions, 2155-2158, 2009.

_____. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação **Revista de Estudos da Linguagem**, Belo Horizonte, v. 20, n. 1, p. 11-27, 2012.

BECKER, M.R. **Análise acústica da produção da nasais bilabiais e alveolares em codas de monossílabos por alunos de inglês**. Dissertação de mestrado não-publicada. Curitiba: Universidade Federal do Paraná, 2007

_____. **Inteligibilidade da língua inglesa sob o paradigma da Língua Franca: Percepção de discursos de falantes de diferentes L1s por brasileiros**. Tese de Doutorado não-publicada. Universidade Federal do Paraná, Curitiba, 2013.

BEST, C. T. A direct realist view of cross-language speech perception. In: STRANGE, W. (Ed.). **Speech perception and linguistic experience: Issues in cross- language research**, p. 171-204, 1995.

BEST, C. T., TYLER, M. D. Nonnative and second-language speech perception: Commonalities and complementarities. In: BOHN, O.-S., MUNRO, M. J. (Eds.). **Second language speech learning: The role of language experience in speech perception and production**. Amsterdam: John Benjamins, p. 13–34, 2007.

BISHOP, J.; KEATING, P. Perception of pitch location within a speaker's range: Fundamental frequency, voice quality and gender. **Journal of the Acoustical Society of America** 132(2), 1100-1112, 2012.

BLANKENSHIP, B. The timing of nonmodal phonation in vowels. **Journal of Phonetics**, 30, 163–191, 2002.

BOERSMA, P.; WEENIK, D. **Praat: doing phonetics by computer**. Version 5.3.82, 2014.

BONGAERTS, T., PLANKEN, B., SCHILS, E. Can late learners attain a native accent in a foreign language? A test of the critical period hypothesis. In SINLGETON, D.; LENGUEL, (eds) Z. **The age factor in second language acquisition**. Clevedon: Multilingual Matters, 30-50, 1995.

BRAWERMAN, A. **Uma análise de erros de estudantes brasileiros de inglês na acentuação de palavras com sufixos**. Dissertação de Mestrado não-publicada. Curitiba: Universidade Federal do Paraná, 2006.

BRAWERMAN-ALBINI, A. **Os efeitos de um treinamento de percepção na aquisição do padrão acentual pré-proparoxítono inglês por estudantes brasileiros**. Tese de doutorado não publicada: Universidade Federal do Paraná, 2012.

BRUYNINCKX, M.; HARMEGNIES, B.; LLISTERRI, J.; e POCH, D. Effects of language change on voice quality. An experimental study of catalan-castilian bilinguals. In **ICPhS 1991. Actes du 12 ème Congrès International de Sciences Phonétiques**. Vol. 2. (pp. 398-401). Aix-en-Provence: Université de Provence, Service des Publications, 1991.

_____. Language-induced voice quality variability in bilinguals. **Journal of Phonetics**, 22, 19-31, 1994.

BYRNE, D.; DILLON, H.; TRAN, K.; ARLINGER, S.; WILBRAHAN, K; COX, R.; HAGER-MAN, B.; HETU, R.; KEI, L.; LUI, C.; KIESSLING, J.; KOTBY, N. M.; NASSER, N. H. A.; WAFAA, NAKANISHI, Y.; OYER, H.; POWELL, R.; STEPHENS, D.; MEREDITH, R.; SIRIMANNA, T., TAVARTKILADZE, G.; FROLENKOV, G. I.; WESTERMAN, S.; LUDVIGSEN, C. "An international comparison of long-term

average speech spectra,” **Journal of the Acoustical Society of America**, 96, 2108–2120. 1994.

CAGLIARI, L.C. Prosódia – algumas funções dos supra-segmentos. **Cadernos de Estudos Lingüísticos**, Campinas (23): p.137-151, 1992.

CAMARGO, Z. **Parâmetros vocais e configurações laríngeas na fonação de indivíduos submetidos às laringectomias parciais verticais**. Dissertação de mestrado não-publicada. São Paulo: Pontifícia Universidade Católica de São Paulo, 1996.

_____. **Análise da qualidade vocal de um grupo de indivíduos disfônicos: uma abordagem interpretativa e integrada de dados de natureza acústica, perceptiva e eletroglotográfica**. Tese de doutorado não-publicada. São Paulo: Pontifícia Universidade Católica de São Paulo, 2002

CAMARGO, Z.; VILARIM, G. S.; CUKIER, S. Parâmetros perceptivo-auditivos e acústicos de longo termo da qualidade vocal de indivíduos disfônicos. **Revista CEFAC**, São Paulo, v. 6, n.2, p. 189-196, 2004.

CAMARGO, Z.; MADUREIRA, S. Voice quality analysis from a phonetic perspective: Voice Profile Analysis Scheme Profile for Brazilian Portuguese (BP VPAS). In **Proceedings of the 4th Conf on Speech Prosody**, v.1, 57±60. Campinas, Brazil, 2008.

CAMARGO, Z.; TSUJI, D.H.; MADUREIRA, S. Analysis of dysphonic voices based on the interpretation of acoustic, physiological and perceptual data. In: Pelethorpe S, Tabain M, editors. **Proceedings of 6th Internacional Seminar on Speech Production**, Sidney. Sidney: Macquaire University. v. 1, p. 31-6, 2003.

CAMARGO, Z.; MADUREIRA, S.; PESSOA, A.; RUSILO, L.C. Voice quality and gender: some insights on correlation between perceptual and acoustic dimensions In: Ma Q, Ding H, Hirst D, editors. **6th International Conference on Speech Prosody**; Shangai. Shangai: Tongji University Press, 2012a. v. 1, p. 115-8, 2012.

CAMARGO, Z.; MADUREIRA, S.; SCHMITZ, J.R. Qualidade de voz e produção de fala em três línguas: um estudo de caso. **Revista Intercâmbio**, v.27, 110-140, 2013.

CARR, P.B.; TRILL, D. Long-term larynx-excitation spectra. **Journal of the Acoustical Society of America**, 36:2033–2040, 1964.

CELCE-MURCIA, M., BRINTON D. M.; GOODWIN, J.. **Teaching pronunciation: A reference for teachers of English to speakers of other languages**. Cambridge: Cambridge University Press, 1996.

CHEN, S. X., BOND, M. H. Two languages, two personalities? Examining language effects on the expression of personality in a bilingual context. **Personality and Social Psychology Bulletin**, 36, 1514-1528, 2010.

CLOPPER, C.G., SMILJANIC, R. Effects of gender and regional dialect on prosodic patterns in American English. **Journal of Phonetics**, 39(2): 237–245, 2011.

COHEN, H.; DOUAIRE, J.; ELSABBAGH, M. The role of prosody in discourse processing. **Brain and Cognition**, 46, p. 73-82, 2001.

CÓRDULA, M.S.M. **Análise fonético-fonológica dos padrões entoacionais do Português Brasileiro e do Inglês Norte-americano no filme Shrek (2001)**. 234f. Tese de Doutorado não-publicada. Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2012.

CRUTTENDEN, A. **Intonation**. Cambridge: Cambridge University Press, 1997.

CRYSTAL, D. Paralinguistics. In BENTHALL, J.; POLHEMUS, H. (Eds) **The Body as a Medium of Expression**. New York: Dutton, 1975.

CUTLER, A.; DAHAN, D.; VAN DONSELAR, W. Prosody in the comprehension of spoken language: A literature review. **Language and Speech**, 40:141-201, 1997.

DE BOT, K.; LOWIE, W.; e VERSPOOR, M. **Second Language Acquisition, an advanced resource book**. London: Routledge, 2005.

DERWING, T. M., MUNRO, M. J., WIEBE, G. E. Evidence in favour of a broad framework for pronunciation instruction. **Language Learning**, 48, 393-410, 1998.

DICANIO, C. T. The phonetics of register in Takhian Thong Chong. **Journal of the International Phonetic Association**, 39, 162–188, 2009.

DOLSON, M. The pitch of speech as a function of linguistic community. **Music Percept.** 11, 321-331, 1994.

ENGELMANN, A. A psicologia da Gestalt e a ciência empírica contemporânea. **Psicologia: Teoria e Pesquisa**, vol . 18 (1), p. 1-16, 2002.

ESLING, J H. Crosslinguistic Aspects of Voice Quality. IN KENT, R.D.; BALL, M.J. (eds). **Voice quality measurement**. Singular Publishing Group: California, 2000.

_____. Voice Quality. **The Encyclopedia of Applied Linguistics**. Wiley Online Library, p.1-7, 2013.

EPSTEIN, M. **Voice quality and prosody in English**. Tese de doutorado não-publicada. University of California, Los Angeles, 2002.

ESPOSITO, C. M. An acoustic and electroglottographic study of White Hmong phonation. **Journal of Phonetics**, 40, 466–476, 2012.

FAIRBANKS, G. **Voice and Articulation Drill Book**. Harper, New York: p. 168, 1940.

FANT, G. **Acoustic Theory of Speech Production**. Mouton: The Hague, 1960.

FANT, G.; LILJENCRAENTS, J.; LIN, Q. "A four-parameter model of glottal flow," **Speech Transm. Lab. Q. Prog. Status Rep.** 4, 1–13, 1985.

FERREIRA, A.P.P. **Pet or petty? diferenças entre palavras CVC e CVCV do inglês por aprendizes brasileiros: uma análise acústica**. Dissertação de mestrado não-publicada. Curitiba: Universidade Federal do Paraná, 2007

FISCHER-JØRGENSEN, E. Phonetic analysis of breathy (murmured) vowels in Gujarati. **Indian Linguistics**, 28, 71–139, 1967.

FLEGE, J. E. Effects of equivalence classification on the production of foreign language speech sounds. In A. James and J. Leather (Eds.), **Sound Patterns in second language acquisition**, Foris Publications, 1987.

_____. Second language speech learning: Theory, findings, and problems. In STRANGE, W. (Ed.) **Speech perception and linguistic experience: Issues in cross-language research**, p.223-277. Timonium, MD: York Press, 1995.

FLEGE J. Age of learning and second-language speech, In BIRDSOING, D (Ed) **Second Language Acquisition and the Critical Period Hypothesis**. Hillsdale, NJ: Lawrence Erlbaum, p. 101-132, 1999.

_____. Interactions between the Native and Second-language Phonetic Systems. In BURMEISTER, P.; PISKE, T.; ROHDE, A. (Eds) **An Integrated View of Language Development: Papers in Honor of Henning Wode**. Trier: Wissenschaftlicher Verlag, p. 217-244, 2002.

_____. Assessing constraints in second language segmental perception and production. In SCHILLER, N.; MEYER A. (Eds.) **Phonetics and Phonology in language comprehension and production: differences and similarities**, 319-355. Berlin: Mouton de Gruyter, 2003.

_____. Language contact in bilingualism: Phonetic system interactions. In COLE, J.; HUALDE, J. (Eds.), **Laboratory Phonology 9**. Berlin: Mouton de Gruyter, p. 353-380, 2007.

FONTES, M. A. S. ; MADUREIRA, S. . Facial and vocal gestures in the expression of emotions. In: S2G From sound to gesture, 2014, Padova. S2G **From sound to gesture**. Padova: **Università degli Studi di Padova**, v. 1. p. 90-91, 2014.

FRØKJAER-JENSEN, B.; PRYTZ, S. Registration of voice quality. **Bruel-Kjaer Technology Review**; 3: 3-17, 1976.

FULLER, B. e HORII, Y. Spectral energy distributions in four types of infant vocalizations. **Journal of Communication Disorders**, 21, 251-261, 1988.

GARELLEK, M. The timing and sequencing of coarticulated non-modal phonation in English and White Hmong. *Journal of Phonetics*, 40 (1): p.152-161, 2012.

GARELLEK, M., e KEATING, P. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. **Journal of the International Phonetic Association**, 41, 185–205, 2011.

GARELLEK, M.; KEATING, P.; ESPOSITO, C.; KREIMAN, J. Voice quality and tone identification in White Hmong. **Journal of the Acoustical Society of America**, 133, 1087–1089, 2013a.

GARELLEK, M.; SALMAN, R.A; KREIMAN, J.; GERRAT, B.R. Perceptual sensitivity to a model of the source spectrum. **Proceedings of Meetings on Acoustics**, Vol. 19, 2013b.

GERRATT, B.R.; KREIMAN, J. Toward a taxonomy of nonmodal phonation. **Journal of Phonetics**, 29, p. 365 – 381, 2001.

GIBBON, F., HARDCASTLE, W. NICOLAIDIS, K. Temporal and spatial aspects of lingual coordination in /kl/ sequences: a cross-linguistic investigation. **Language and Speech** 36, 261–77, 1993.

GOBERMAN, A. M., JONHNSON, S., CANNIZZARO, M., e ROBB, M. The effect of infant positioning on infant cries: Implications for sudden infant death syndrome. **International Journal of Pediatric Otorhinolaryngology**, 72, 153-165, 2008.

GOBL, C.; NI CHASAIDE, A. Voice source variation in the vowel as a function of consonantal context. In HARDCASTLE, W. J.; HEWLETT, N. (Eds.), **Coarticulation: Theory, data and techniques**. Cambridge: Cambridge University Press, p. 122–143, 1999.

GOGGIN, J., THOMPSON, C., StTRUBE, G., SIMENTAL, L. The role of language familiarity in voice identification. **Memory Cognition**, 19 (5), 448–458, 1991.

GOMES, M. L. C. **A produção de palavras do inglês com o morfema –ed por falantes brasileiros: uma visão dinâmica**. Tese de Doutorado. Universidade Federal do Paraná, Curitiba, 2009.

GORDON, M. The phonetics and phonology of nonmodal vowels: a cross-linguistic perspective. **Berkeley Linguistics Society**, 24, 93-105, 1998.

GORDON, M.; LADEFOGED, P. Phonation types: a cross-linguistic overview. **Journal of Phonetics**, 29:383-406, 2001.

GROSJEAN, F. **Studying bilinguals**. Oxford: OUP, 2008.

GUT, U. **Non-native speech: A corpus-based analysis of phonological and phonetic properties of L2 English and German**. Frankfurt, Germany: Peter Lang, 2009.

HAMMARBERG, B., FRITZELL, B., GAUFFIN, J., SUNDBERG, J.; WEDIN, L. "Perceptual and acoustic correlates of abnormal voice qualities", **Acta Otolaryngol** 90, 441–451, 1980.

HANSON, H. M. **Glottal characteristics of female speakers**. Tese de doutorado não-publicada. MIT: 1995.

_____. Glottal characteristics of female speakers: Acoustic correlates. **Journal of the Acoustical Society of America** 101(1), 466–481, 1997.

HANSON, H. M.; CHUANG, E. S. Chuang. Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. **Journal of the Acoustical Society of America** 106, 1064–1077, 1999.

HANSON, H. M.; STEVENS, K.N.; KUO, H-K. J.; CEHN, M.Y.; SLIFKA, J. Towards models of phonation. **Journal of Phonetics** 29(4), 451–480, 2001.

HARMEGNIES, B.; LANDERCY, A. Language features in the long-term average spectrum. **Revue de Phonétique Appliquée**, 73-75: 69-79, 1985.

HARMEGNIES, B.; BRUYNINCKX, M.; LLISTERRI, J.; e POCH, D. Effects of language change on voice quality: An experimental contribution to the study of the Catalan-Castilian case. In **Eurospeech 1989: Proceedings of the 1st European Conference on Speech Communication and Technology**, 489-492, 1989.

_____. Effects of language change in voice quality in bilingual speakers: Corpus content effects. In **Eurospeech 1991. Proceedings of the 2nd European conference on speech communication and technology**, Vol. 1, p.165-8. Genova, Italy, 1991.

HILLENBRAND, J.M.; CLEVELAND, R.A.; and ERICKSON, R.L. "Acoustic correlates of breathy vocal quality," **Journal of Speech and Hearing Research**, 37, 769-778, 1994.

HILLENBRAND, J.; HOUDE, R.A. Acoustic correlates of breathy voice quality: dysphonic voices and continuous speech. **Journal of Speech, Language, and Hearing**, 39, 298-310, 1996.

HIRST, D.J.; DI CRITO, A. (eds). **Intonation Systems. A survey of Twenty Languages**. Cambridge: Cambridge University Press, 1998.

HOLLIEN, H.; MICHAEL, J.F. Vocal fry as a phonational register. **Journal of Speech and Hearing Research**, 11, p. 600-604, 1968.

HOLMBERG, E. B.; HILLMAN, R. E.; PERKELL, J. S. Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal, and high pitch. **Journal of Voice**, 3: p. 294-305, 1989.

HOLMBERG, E. B.; HILLMAN, R. E.; PERKELL, J. S.; GUIOD, P. C.; GOLDMAN, S. L. Comparisons among aerodynamic, electroglottographic and acoustic spectral measures of female voice. **Journal of Speech and Hearing Research** 38:1212-1223, 1995.

HUANG, B., JUN, S. Age effect on the acquisition of second language prosody. In **Proceedings of the 33rd Boston University Conference on Language Development**, 2009.

ISELI, M.; ALAWAN, A. An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation. **Proceedings of the ICASSP**, Montreal, May 2004, p. 669–672, 2004.

ISELI, M., SHUE, Y.-L.; EPSTEIN, M.; KEATING P.; KREIMAN, J.; ALWAN A. Voice source correlates of prosodic features in American English: A pilot study. **Proceedings of Interspeech 2006**, Pittsburgh, October 2006.

ISELI, M., SHUE, Y.-L.; ALWAN, A. Age, sex, and vowel dependencies of acoustic measures related to the voice source, **Journal of the Acoustical Society of America**, 121, 2283–2295, 2007.

JILKA, M. **The contribution of intonation to the perception of foreign accent**. Unpublished doctoral dissertation. University of Stuttgart, Germany, 2000.

JONHSON, K. **Acoustic and Auditory Phonetics**. 2nd Edition. Oxford: Blackwell, 2012.

JONES, R.H.; EVANS, S. Teaching pronunciation through voice quality. **ELT Journal** 49: 3 244- 251, 1995.

KARPINSKI, M. The Boundaries of Language: Dealing with Paralinguistic Features. **Lingua Posnaniensis**, vol. LIV (2), p. 37-54, 2012

KEATING, P.A.; ESPOSITO E, C. Linguistic Voice Quality. **UCLA Working Papers in Phonetics**, nº 105, p.85-91, 2007.

KEATING, P.; ESPOSITO, C.; GARELLEK, M., KHAN, S.; KUANG, J. Phonation contrasts across languages. Poster presented at **LabPhon12** in Albuquerque NM, 2010.

KEATING, P. A.; KUO, G. Comparison of speaking fundamental frequency in English and Mandarin. **Journal of the Acoustic Society of America**, 132 (2), P. 1050 – 1060, 2012.

KELLER, E. The Analysis of Voice Quality in Speech Processing. (Manuscript for the Tutorial Research Workshop **Nonlinear Speech Processing: Algorithms and Analysis**, Vietri, Italy, 2004.

KENT, R.D.; BALL, M.J. (eds). **Voice quality measurement**. Singular Publishing Group: California, 2000.

KHAN, S.D. The phonetics of contrastive phonation in Gujarati. **Journal of Phonetics**, v. 40, issue 6, 780-795, 2012.

KIRK, P. L., LADEFOGED, J.; LADEFOGED, P. Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec. In **American Indian linguistics and ethnography in honor of Laurence C. Thompson** (A. Mattina e T. Montler, editors), Missoula, MT: University of Montana Press, 1993.

KITZING, P. LTAS criteria pertinent to the measurement of voice quality. **Journal of Phonetics**, v. 14, p. 477- 482, 1986.

KLUGE, D. C. **Perception and production of English syllable-final nasals by Brazilian learners**. Dissertação de Mestrado não-publicada. Florianópolis: Universidade Federal de Santa Catarina, 2004.

_____. **Brazilian EFL learners' identification of wordfinal /m-n/: Native/Nonnative realization and effects of visual cues**. Tese de Doutorado não-publicada. Florianópolis: Universidade Federal de Santa Catarina, 2009.

KOERICH, R. D. **Perception and production of word-final vowel epenthesis by Brazilian EFL students**. Tese de Doutorado não-publicada. Florianópolis: Universidade Federal de Santa Catarina, 2002.

_____. Perception and production of vowel paragoge by Brazilian EFL students. In BAPTISTA, B.O.; WATKINS, M. A.; (Eds.), **English with a Latin beat: studies in Portuguese/Spanish – English interphonology**, p. 91-104, Amsterdam: John Benjamins, 2006.

KREIMAN, J. Perception of sentence and paragraph boundaries in natural conversation. **Journal of Phonetics**, 10, p. 163-175, 1982.

_____. **Human Memory for Unfamiliar Voices**. Tese de Doutorado não-publicada. Chicago-IL, University of Chicago, 1987.

KREIMAN, J.; SIDTIS, D.V.; GERRATT, B. Defining and measuring voice quality. **From Sound to Sense**: MIT, 2004.

KREIMAN, J.; VANLACKER-SIDTIS, D.; GERRATT, B. R. Perception of voice quality. IN PISONI, D.B.; REMEZ, R.E (eds) **Handbook of Speech Perception**. Blackwell, Oxford, pp. 338–362, 2005.

KREIMAN, J.; GERRATT, B. R.; ANTONANZAS-BARROSO, N. Measures of glottal source spectrum. **Journal of Speech and Hearing Research**, 50, 595–610, 2007.

KREIMAN, J.; ANTONANZAS-BARROSO, N.; GERRATT, B. R. Integrated software for analysis and synthesis of voice quality. **Behavior Research Methods**, 42, 2010.

KREIMAN, J.; SIDTIS, D. **Foundations of Voice Studies: An Inter-disciplinary Approach to Voice Production and Perception**. Wiley-Black-well, Walden, MA, pp. 1–518, 2011.

KREIMAN, J.; SHUE, Y.-L.; CHEN, G.; ISELI, M.; GERRATT, B. R.; NEUBAUER, J.; e ALWAN, A. Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. **Journal of the Acoustical Society of America**, 132, 2625–2632, 2012.

KREIMAN, J.; GERRATT, B.R.; GARELLEK, M.; SALMAN, R.; ZHANG, Z. Toward a unified theory of voice production and perception. **Loquens**, 1 (1), p. 1-9, 2014.

KREIMAN, J.; GARELLEK, M.; SALMAN, R. A.; GERRATT, B. R. Perceptual sensitivity to a model of the voice source spectrum. Manuscript in preparation, 2014b.

LADEFOGED, P. **Preliminaries to linguistic phonetics**. Chicago: University of Chicago Press, 1971.

_____. The linguistic use of different phonation types. In D. Bless e J. Abbs (Eds.), *Vocal fold physiology: Contemporary research and clinical issues*. San Diego: College Hill Press, p. 351-360, 1983.

LADEFOGED, P.; MADDIESON, I. **Sounds of the World's Languages**. Blackwells, Oxford, 1996.

LAVER, J. **The phonetic description of voice quality**. Cambridge: Cambridge University Press, 1980.

_____. **The gift of speech: papers in the analysis of speech and voice**. Edinburgh: Edinburgh University Press, 1991.

_____. The phonetic evaluation of voice quality. In: KENT, R. D.; BALL, M.J. (eds). **Voice quality measurement**. San Diego: Singular Publishing, p. 37-48, 2000.

LAVER, J.; WIRS, S.; MACKENZIE-BECK, J.; HILLER, S.M. A perceptual protocol for the analysis of vocal profiles. In: LAVER, J. (Ed.). **The gift of speech: papers in the analysis of speech and voice**. Edinburgh: Edinburgh University Press; p 265-280, 1991.

LAVER, J. e TRUDGILL, P. Phonetic and linguistic markers in speech. In: K.R. Scherer e H. Giles (eds.). **Social markers in speech**. Cambridge: Cambridge University Press, p.1-32, 1979.

LEHISTE, I. **Suprasegmentals**. MIT Press, Cambridge, Mass, 1970.

LENGERIS, A. Prosody and Second Language Teaching: Lessons from L2 Speech Perception and Production Research. IN: ROMERO-TRILLO, J. (ED) **Pragmatics and Prosody in English Language Teaching**. Springer: Madrid, p.25-40, 2012.

LENNEBERG, E.H. **Biological foundations of language**. New York: Wiley, 1967.

LINDENBAUM, L. **Long-time Average Spectrum in Individuals with Parkinson Disease**. (Electronic Thesis or Dissertation). University of Ohio, 2012.

LIMA, M. F. B. **Sintomas vocais, alteração de qualidade vocal e laríngea em professores**. Dissertação de mestrado não-publicada. São Paulo: Pontifícia Universidade Católica de São Paulo, 2008.

LIMA-SILVA, M. F. B. **Avaliação da qualidade vocal com motivação fonética: análise integrada de dados de percepção e acústica**. Tese de doutorado não-publicada. São Paulo: Pontifícia Universidade Católica de São Paulo, 2012.

LÖFQVIST, A.; MCGOWAN, R. S. Influence of consonantal environment on voice source aerodynamics. **Journal of Phonetics**, 20, p. 93–110, 1992.

LOTTO, A.J.; HOLT, L.L.; KLUENDER, K.R. Effect of voice quality on perceived height of English vowels. **Phonetica**, 54, p.76-93, 1997.

MACK, M. The phonetic systems of bilinguals. In BANICH, M.; MACK, M. (Eds.), **Mind, brain, and language: Multidisciplinary perspectives**, p. 309-349, 2003.

MACKAY, I., MEADOR, D., e FLEGE, J. The identification of English consonants by native speakers of Italian. **Phonetica**, 58, 103-125, 2001.

MACKENZIE-BECK, J. Perceptual analysis of voice quality: the place of vocal profile analysis. In: HARDCASTLE, W.J.; MACKENZIE-BECK, J. **A figure of speech: a festschrift for John Laver**. Mahwah, NJ: Lawrence Erlbaum Associates; p. 285-322, 2005.

MADUREIRA, S. A expressão de atitudes e emoções na fala. In: Leny Kirillos. (Org.). **Expressividade**. São Paulo: Revinter, p. 15-25. 2004.

MADUREIRA, S.; CAMARGO, Z. A. Exploring sound symbolism in the investigation of speech expressivity. In: ISCA, 2010, Athens. **Exploring sound symbolism in the investigation of speech expressivity**. Athens: Edited by Antonis Botinis, v. 01. p. 105-108, 2010.

MARTA, E. S. **Avaliação e percepção da prosódia na expressão de modalidades por falantes de inglês como L1 e aprendizes de inglês como L2**.

Tese de Doutorado não publicada. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2011

MARYN, Y.; ROY, N.; DE BODT, M.; VAN CAWENBERGE, P.; CORTHALS, P. Acoustic measurement of overall quality: a meta-analysis. **Journal of the Acoustic Society of America**, 126 (5), p. 2619-2634, 2009, 2009.

MAUAD, S.A. **Questões de prosódia: uma investigação, com o apoio de instrumentais de análise fonético-acústica, dos padrões entoacionais de falantes bilíngues brasileiros e norte-americanos**. Dissertação de mestrado não-publicada. São Paulo: Pontifícia Universidade Católica de São Paulo, 2007.

MENNEN, I. Bi-directional interference in the intonation of Dutch speakers of Greek. **Journal of Phonetics**, 32, 543-563, 2004.

_____. Phonetic and phonological influences in non-native intonation: an overview for language teachers. **QMUC Speech Science Research Centre Working Papers**, WP-9, 2006.

MENNEN I., LEEUW, E., SCOBIE, J. M., SCHAEFFLER, F. e SCHAEFFLER, S. Measuring language-specific phonetic settings. **Second Language Research**, 26 (1), 13-41, 2010.

MENNEN, I.; LEEUW, E. Beyond Segments. **Studies in Second Language Acquisition**, 36, p. 183-194, 2014.

MENNEN, I.; SCHAEFFLER, F.; DICKIE, C. Second Language Acquisition of Pitch Range in German learners of English. **Studies in Second Language Acquisition**, 36, p. 303-329, 2014.

MOYER, A. Age, accent and experience in second language acquisition. Clevedon: **Multilingual Matters**, 2004.

_____. Nonsegmental factors in foreign accent. **Studies in Second Language Acquisition**, 17 (1), p. 17-33, 1995.

MUNRO, M. J. The effects of noise on the intelligibility of foreign-accented speech. **Studies in Second Language Acquisition** 20: 139–154, 1998.

MUNRO, M. J.; DERWING, T. M. Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. **Language Learning**, 45:1, p. 73-97, 1995.

_____. Modelling perceptions of the comprehensibility and accentedness of L2 speech: The role of speaking rate. **Studies in Second Language Acquisition**, 23, 451-468, 2001.

NI CHASAIDE, A.; GOBL, C. Voice source variation. In HARDCASTLE, W. J.; LAVER, J. (Eds.), **The handbook of phonetic sciences**. Oxford, UK: Blackwell, p. 427–461, 1997.

NG, M. L.; CHEN, Y.; CHAN, E. Y. K. Differences in Vocal Characteristics Between Cantonese and English Produced by Proficient Cantonese-English Bilingual Speakers—A Long-Term Average Spectral Analysis. **Journal of Voice**, Vol. 26, No. 4, p. e171-e176, 2012.

NOBRE-OLIVEIRA, D. **The effect of perceptual training on the learning of English vowels by Brazilian Portuguese speakers**. Tese de Doutorado não-publicada. Florianópolis: Universidade Federal de Santa Catarina, 2007.

OLIVEIRA, L.R. **Correlatos acústicos e perceptivos da postura de língua na produção de fala de respiradores orais**. Dissertação de mestrado não-publicada. São Paulo: Pontifícia Universidade Católica de São Paulo, 2011.

PEPIOT, E. Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers. **Proceedings of the 7th International Conference on Speech Prosody** - Dublin, 305-309, 2014.

PESSOA, A.; NOVAES, B.C.A.C.; MADUREIRA, S.; CAMARGO, Z. Perceptual and acoustic correlates of a speech in a bilateral cochlear implant user. In: **6th International Conference on Speech Prosody**; 2012; Shangai. Shangai: Tongji University Press; v. 1. p. 51-4, 2012.

PIKE, K. **The Intonation of American English**. Ann Arbor, MI: University of Michigan Press, 1945.

PITTAM, J. The long term spectral measurement of voice quality as a social and personality marker : a review. In **Tadjen**, Kris, 1-12, 1987.

PODESVA, R.J. Gender and the social meaning of non-modal phonation types. **Proceedings of the Berkeley Linguistics Society**, v.37, 427-448, 2013.

RAUBER, A. S. **Perception and production of English vowels by Brazilian EFL speakers**. Tese de Doutorado não-publicada. Florianópolis: Universidade Federal de Santa Catarina, 2006.

RAUBER, A., RATO, A., KLUGE, D.C.; SANTOS, G. R.; FIGUEIREDO, M. **TP – Teste de Percepção**. Versão. 3.1, 2012.

ROCCA P. D. A. O impacto da tecnologia de fala no ensino da entoação de uma língua estrangeira. **Revista Intercâmbio**, v. XXIII: 1-24, São Paulo: LAEL/PUCSP, 2011.

RUSILO, L. C. ; CAMARGO, Z. A. ; MADUREIRA, Sandra. The validity of some acoustic measures to predict voice quality settings: trends between acoustic and perceptual correlates of voice quality. In: **4TH Isca Tutorial and Research**

Workshop on Experimental Linguistics, Paris, p. 115-118, 2011.

SALMAN, R. A.; STORY, B. H. Relation of structural and vibratory kinematics of the vocal folds to two acoustic measures of breathy voice based on computational modeling. **Journal of Speech, Language, e Hearing Research**, **54**, 1267–1283, 2011.

SALMAN, R. A.; STORY, B. H.; BUNTON, K. Relation of perceived breathiness to laryngeal kinematics and acoustic measures based on computational modeling. **Journal of Speech, Language, e Hearing Research**, **56**, 1209–1223, 2013.

SAMELY, U. **Kedang (Eastern Indonesia), some aspects of its grammar**. Hamburg: Helmut Buske Verlag., 1991.

SHATTUCK-HUFNAGEL, S.; TURK, A. A prosody tutorial for investigators of auditory sentence processing. **Journal of Psycholinguistic Research** 25(2). 193-247, 1996.

SILVEIRA, R. **The influence of pronunciation instruction on the perception and production of English word-final consonants**. Tese de Doutorado não-publicada. Florianópolis: Universidade Federal de Santa Catarina, 2004.

SHUE Y.-L., KEATING, P., VICENIK, C., YU, K. "VoiceSauce: A program for voice analysis," in **Proc. 17th ICPHS Hong Kong**, pp. 1846–1849, 2011.

SUNDBERG, J.; NORDENBERG, M., "Effects of vocal loudness variation on spectrum balance as reflected by the alpha measure of long-term-average spectra of speech" *J. Acoust. Soc. Am.*, **120**, 453–457 (2006).

TAKIMA, K.; PORT, R.; DALBY, J. Effects of temporal correction on intelligibility of foreign-accented English. **Journal of Phonetics**, **25**, 1–24, 1997.

TODAKA, Y. A preliminary study of voice quality differences between Japanese and American English: Some pedagogical suggestions. **JALT journal**, **17**(2), 261-268, 1995.

TROFIMOVICH, P., BAKER, W. Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. **Studies in Second Language Acquisition**, **28**, 1-30, 2006.

TROFIMOVICH, P.; BAKER, W. Learning prosody and fluency characteristics of second-language speech: Child learners' acquisition of five suprasegmentals. **Applied Psycholinguistics**, **28**, 251-276, 2007.

VIOLA, I. C. ; MADUREIRA, S. A qualidade vocal como elemento expressivo. In: Aragão, M. S.S.; Silveira, R. C.. (Org.). **Diversidade de estudos em Fonética e Fonologia no Brasil**. João Pessoa - PB: ANPOLL 2007, p. 78-86, 2007.

WESTER, M. Talker Discrimination Across Languages. **Speech Communication**, 54, p. 781-790, 2012.

WESTER, M., DINES, J., GIBSON, M., LIANG, H., WU, Y.-J., SAHEER, L., KING, S., OURA, K., GARNER, P., BYRNE, W., GUAN, Y., HIRSIMA, T., KARHILA, R., KURIMO, M., SHANNON, M., SHIOTA, S., TIAN, J., TOKUDA, K., YAMAGISHI, J. Speaker adaptation and the evaluation of speaker similarity in the EMIME speech-to-speech translation project. In: **Proceedings of SSW7**, 2010.

WILSON, I. **Articulatory settings of French and English monolingual and bilingual speakers**. Unpublished PhD thesis, University of British Columbia, 2006

WINTERS, S., LEVI, S., PISONI, D. Identification and discrimination of bilingual talkers across languages. **Journal of the Acoustical Society of America**. 123, 4524-4538, 2008.

ZHANG, Z.; KREIMAN, J.; GERRAT, B. R.; GARELLEK, M. Acoustic and perceptual effects of changes in body layer stiffness in symmetric and asymmetric vocal fold models. **Journal of the Acoustical Society of America**, 133, 453–462, 2013.

ZIMMER, M. C. ; ALVES, U. K. . O impacto do bi/multilinguismo sobre o potencial criativo em sala de aula: uma análise via teoria dos sistemas dinâmicos. **Revista FAEEBA**, v. 23, p. 77-90, 2014.

LISTA DE ANEXOS

ANEXO A-	TEXTOS UTILIZADOS NO EXPERIMENTO PILOTO	157
ANEXO B-	LISTAS DE SENTENÇAS-VEÍCULO DO EXPERIMENTO DE PRODUÇÃO	158
ANEXO C-	TEXTOS UTILIZADOS NO EXPERIMENTO DE PRODUÇÃO	162
ANEXO D-	PERGUNTAS UTILIZADAS NO EXPERIMENTO DE PRODUÇÃO	163
ANEXO E-	ROTEIRO DE ENTREVISTA PARA OS PARTICIPANTES DO EXPERIMENTO DE PRODUÇÃO	164
ANEXO F-	TERMO DE CONSENTIMENTO DO EXPERIMENTO DE PRODUÇÃO	165
ANEXO G-	ROTEIRO DE ENTREVISTA PARA OS PARTICIPANTES DO EXPERIMENTO DE PERCEPÇÃO	167

ANEXO A

A Cigarra e a formiga

Em um campo, num dia de verão, uma cigarra saltitava e cantava alegremente com todo seu coração. Uma formiga passou levando consigo uma espiga de milho com grande esforço, e ia em direção ao ninho.

"Por que não vem conversar comigo", disse a Cigarra, "em vez de trabalhar dessa maneira?"

"Estou guardando comida para o inverno", disse a Formiga, "e recomendo que você faça o mesmo."

"Por que se preocupar com o inverno?" disse a Cigarra, "temos alimentos em abundância no presente".

Mas a Formiga continuou o seu caminho e seu trabalho. Quando o inverno chegou, a Cigarra não tinha comida e se viu morrendo de fome, enquanto via as formigas diariamente distribuindo milho e grãos dos estoques que haviam coletado no verão.

Então a Cigarra soube: É melhor se preparar para os dias de necessidade.

The Ant and the Grasshopper

In a Field, one summer's Day, a Grasshopper was hopping about, chirping and singing to its heart's content. An ant passed by, bearing along with great toil, an ear of corn it was taking to the nest.

"Why not come and chat with me", said the Grasshopper, "instead of toiling and moiling in that way?"

"I'm helping to lay up food for the winter", said the Ant, "and recommend you to do the same".

"Why bother about winter?", said the Grasshopper, "we have got plenty of food at present".

But the Ant went on its way and continued its toil. When the winter came, the Grasshopper had no food and found itself dying of hunger, while it saw the ants distributing everyday corn and grain from the stores they had collected in the summer.

Then, the Grasshopper know: It is best to prepare for the days of necessity.

ANEXO B

Frases-veículo – Lista 1 - PB

- | | |
|------------------------|------------------------|
| 1. Diga paba pra ela. | 17. Diga papa pra ela. |
| 2. Diga tapa pra ela. | 18. Diga gapo pra ela. |
| 3. Diga lapa pra ela. | 19. Diga vaca pra ela. |
| 4. Diga taco pra ela. | 20. Diga gaba pra ela. |
| 5. Diga capa pra ela. | 21. Diga gaco pra ela. |
| 6. Diga vaga pra ela. | 22. Diga taga pra ela. |
| 7. Diga caco pra ela. | 23. Diga gago pra ela. |
| 8. Diga paca pra ela. | 24. Diga laba pra ela. |
| 9. Diga caga pra ela. | 25. Diga baco pra ela. |
| 10. Diga taba pra ela. | 26. Diga laca pra ela. |
| 11. Diga bapa pra ela. | 27. Diga pago pra ela. |
| 12. Diga cabo pra ela. | 28. Diga daca pra ela. |
| 13. Diga baba pra ela. | 29. Diga lago pra ela. |
| 14. Diga dapa pra ela. | 30. Diga vapa pra ela. |
| 15. Diga vaba pra ela. | 31. Diga dabo pra ela. |
| 16. Diga dago pra ela. | 32. Diga baga pra ela. |

Frases-veículo – Lista 2 - PB

- | | |
|-----------------------|------------------------|
| 1. Diga baco pra ela. | 7. Diga bapa pra ela. |
| 2. Diga paca pra ela. | 8. Diga caco pra ela. |
| 3. Diga caga pra ela. | 9. Diga baba pra ela. |
| 4. Diga tapa pra ela. | 10. Diga dapa pra ela. |
| 5. Diga vaga pra ela. | 11. Diga vaba pra ela. |
| 6. Diga taba pra ela. | 12. Diga dago pra ela. |
| | 13. Diga cabo pra ela. |

- | | |
|------------------------|------------------------|
| 14. Diga papa pra ela. | 24. Diga laca pra ela. |
| 15. Diga gapo pra ela. | 25. Diga capa pra ela. |
| 16. Diga vaca pra ela. | 26. Diga pago pra ela. |
| 17. Diga gaba pra ela. | 27. Diga daca pra ela. |
| 18. Diga taco pra ela. | 28. Diga lago pra ela. |
| 19. Diga paba pra ela. | 29. Diga vapa pra ela. |
| 20. Diga gaco pra ela. | 30. Diga lapa pra ela. |
| 21. Diga taga pra ela. | 31. Diga dabo pra ela. |
| 22. Diga gago pra ela. | 32. Diga бага pra ela. |
| 23. Diga laba pra ela. | |

Frases-veículo – Lista 3 - PB

- | | |
|------------------------|------------------------|
| 1. Diga dapa pra ela. | 17. Diga bapa pra ela. |
| 2. Diga baco pra ela. | 18. Diga gago pra ela. |
| 3. Diga cabo pra ela. | 19. Diga laba pra ela. |
| 4. Diga dago pra ela. | 20. Diga caga pra ela. |
| 5. Diga papa pra ela. | 21. Diga laca pra ela. |
| 6. Diga caco pra ela. | 22. Diga capa pra ela. |
| 7. Diga gapo pra ela. | 23. Diga pago pra ela. |
| 8. Diga vaca pra ela. | 24. Diga daca pra ela. |
| 9. Diga gaba pra ela. | 25. Diga lago pra ela. |
| 10. Diga taco pra ela. | 26. Diga vapa pra ela. |
| 11. Diga vaba pra ela. | 27. Diga baba pra ela. |
| 12. Diga paba pra ela. | 28. Diga lapa pra ela. |
| 13. Diga tapa pra ela. | 29. Diga vaga pra ela. |
| 14. Diga gaco pra ela. | 30. Diga dabo pra ela. |
| 15. Diga paca pra ela. | 31. Diga taga pra ela. |
| 16. Diga taba pra ela. | 32. Diga бага pra ela. |

Frases-veículo – Lista 1 - IN

- | | |
|------------------------|------------------------|
| 1. Say babble to her. | 17. Say gaga to her. |
| 2. Say cabbie to her. | 18. Say baggie to her. |
| 3. Say backer to her. | 19. Say lappy to her. |
| 4. Say tabby to her. | 20. Say kappa to her. |
| 5. Say dapper to her. | 21. Say labble to her. |
| 6. Say tapas to her. | 22. Say bappy to her. |
| 7. Say dabble to her. | 23. Say lackey to her. |
| 8. Say caggy to her. | 24. Say paggy to her. |
| 9. Say dacker to her. | 25. Say lagger to her. |
| 10. Say pabber to her. | 26. Say vapid to her. |
| 11. Say dagger to her. | 27. Say gabble to her. |
| 12. Say gappy to her. | 28. Say vabble to her. |
| 13. Say tacos to her. | 29. Say khaki to her. |
| 14. Say pappy to her. | 30. Say vacky to her. |
| 15. Say gackey to her. | 31. Say packer to her. |
| 16. Say tagger to her. | 32. Say vaggy to her. |

Frases-veículo – Lista 2 - IN

- | | |
|-----------------------|------------------------|
| 1. Say cabbie to her. | 10. Say backer to her. |
| 2. Say gackey to her. | 11. Say pabber to her. |
| 3. Say dapper to her. | 12. Say dacker to her. |
| 4. Say tacos to her. | 13. Say kappa to her. |
| 5. Say tagger to her. | 14. Say labble to her. |
| 6. Say gaga to her. | 15. Say bappy to her. |
| 7. Say pappy to her. | 16. Say lackey to her. |
| 8. Say baggie to her. | 17. Say babble to her. |
| 9. Say lappy to her. | 18. Say tapas to her. |

- | | |
|------------------------|------------------------|
| 19. Say paggy to her. | 26. Say dabble to her. |
| 20. Say lagger to her. | 27. Say vacky to her. |
| 21. Say gappy to her. | 28. Say vabble to her |
| 22. Say vapid to her. | 29. Say packer to her. |
| 23. Say gabble to her. | 30. Say caggy to her. |
| 24. Say tabby to her. | 31. Say dagger to her |
| 25. Say khaki to her. | 32. Say vaggy to her. |

Frases-veículo – Lista 3 – IN

- | | |
|------------------------|------------------------|
| 1. Say dapper to her. | 17. Say vabble to her. |
| 2. Say kappa to her. | 18. Say khaki to her. |
| 3. Say tacos to her. | 19. Say dabble to her. |
| 4. Say labble to her. | 20. Say gackey to her. |
| 5. Say bappy to her. | 21. Say baggie to her. |
| 6. Say lackey to her. | 22. Say vacky to her. |
| 7. Say babble to her. | 23. Say backer to her. |
| 8. Say tapas to her. | 24. Say tabby to her. |
| 9. Say tagger to her. | 25. Say packer to her. |
| 10. Say dacker to her. | 26. Say caggy to her. |
| 11. Say pabber to her. | 27. Say lappy to her. |
| 12. Say paggy to her | 28. Say vaggy to her. |
| 13. Say gappy to her. | 29. Say dagger to her. |
| 14. Say cabbie to her. | 30. Say pappy to her. |
| 15. Say vapid to her. | 31. Say lagger to her. |
| 16. Say gabble to her. | 32. Say gaga to her. |

ANEXO C

Primeiro parágrafo do “Rainbow Passage”

"When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow."

Versão em PB

“Quando a luz do sol atinge os pingos de chuva no ar, eles agem como prisma e formam um arco-íris. O arco-íris é uma divisão da luz branca em várias outras cores. Estas assumem o formato de um longo arco arredondado, com sua trajetória na parte de cima, e suas duas extremidades parecem ir além do horizonte. De acordo com a lenda, há um pote de ouro em uma das suas extremidades. As pessoas procuram, mas nunca o encontram. Quando um homem procura por algo além do seu alcance, seus amigos dizem que ele está a procura do pote de ouro no fim do arco-íris.”

ANEXO D

Perguntas em PB

1. Onde você nasceu? Me conte um pouco sobre a cidade. Você pode falar sobre: tamanho da cidade, clima, população, ou outras coisas que desejar.
2. Há algo que você não gosta sobre Los Angeles/Estados Unidos? Você pode falar sobre: cultura, comida, estilo de vida, pessoas, ou outras coisas que desejar.
3. Do que você mais sente falta do Brasil? Você pode falar sobre: família, amigos, clima, comida, ou outras coisas que desejar.

Perguntas em IN

1. Where were you born? Tell me a little about the city. You can talk about the size of the city, its weather and population, or other things you wish to speak about.
2. Is there something you don't like or hate about Los Angeles/ the USA? You can talk about culture, food, lifestyle, people, or other things you wish to speak about.
3. What do you miss most about Brazil? You can talk about family, friends, food, weather, or other things you wish to speak about.

ANEXO E**ENTREVISTA**

Participante # _____

Sexo: _____

Data de Nascimento _____

Local de Nascimento/Onde cresceu _____

Há quanto tempo está nos EUA/Los Angeles?

Com que idade começou a estudar inglês?

Por quanto tempo estudou inglês? Onde estudou?

Por quanto tempo mora/morou em país que fala língua inglesa?

Como é o uso do PB e do IN no seu dia-a-dia (em % aproximada)?

Fala alguma outra língua estrangeira?

Outros

comentários:

ANEXO F

UNIVERSITY OF CALIFORNIA LOS ANGELES STUDY INFORMATION SHEET

A Cross-linguistic Study on Voice Quality

Jody Kreiman, PhD, from the Department of Head and Neck Surgery at the University of California, Los Angeles (UCLA) and Ana Paula Petriu Ferreira Engelbert, Msc, Visiting Graduate Researcher at UCLA, are conducting a research study.

You were selected as a possible participant in this study because you speak Brazilian Portuguese and/or English fluently and you do not have voice or hearing disorders. Your participation in this research study is voluntary.

1) Why is this study being done?

Our main objective is to analyze the speech of Brazilian participants who speak Portuguese and English fluently. To be able to analyze their speech in English, we will also record speech from native speakers of English, who will be part of the control group.

2) What will happen if I take part in this research study?

If you volunteer to participate in this study, the researcher will ask you to do the following:

Answer some questions about your personal data (age, place of birth, age of learning English, age of arrival in the US, use of Portuguese and English, complaints of voice or hearing disorders).

Be recorded speaking Brazilian Portuguese and English in a soundproof room in the Voice Perception Laboratory at UCLA. Alternatively, you will be recorded in a quiet room with the use of a laptop computer if coming to the Voice Perception Lab is not practical. You will be asked to read sentences or short texts, and to answer simple questions about your daily life.

3) How long will I be in the research study?

Participation will take a total of about 30 to 40 minutes.

4) Are there any potential risks or discomforts that I can expect from this study?

There are no anticipated risks or discomforts.

5) Are there any potential benefits if I participate?

You will not directly benefit from your participation in the research.

The results of the research may shed some light on the different ways people use their voices to speak foreign languages, which may help improve language teaching.

6) Will I be paid for participating?

You will receive \$20.00 in cash.

7) Will information about me and my participation be kept confidential?

Any information that is obtained in connection with this study and that can identify you will remain confidential. It will be disclosed only with your permission or as

required by law. Confidentiality will be maintained by means of coding names in numbers (for example, Subject 3). All your data will be safely kept in the Voice Perception Laboratory and the only people who will have access to it are the investigators directly involved in the study.

8) What are my rights if I take part in this study?

You can choose whether or not you want to be in this study, and you may withdraw your consent and discontinue participation at any time.

Whatever decision you make, there will be no penalty to you, and no loss of benefits to which you were otherwise entitled.

You may refuse to answer any questions that you do not want to answer and still remain in the study.

9) Who can I contact if I have questions about this study?

a. The research team:

If you have any questions, comments or concerns about the research, you can talk to one of the researchers. Please contact:

Jody Kreiman, PhD, and Ana Paula Petriu Ferreira Engelbert at (310) 825-0736.
UCLA Rehab Center

1000 Veteran Ave. UCLA School of Medicine. Los Angeles, CA. 90095.

b. UCLA Office of the Human Research Protection Program (OHRPP):

If you have questions about your rights while taking part in this study, or you have concerns or suggestions and you want to talk to someone other than the researchers about the study, please call the OHRPP at (310) 825-7122 or write to:

UCLA Office of the Human Research Protection Program

11000 Kinross Avenue, Suite 211, Box 951694

Los Angeles, CA 90095-1694

You will be given a copy of this information to keep for your records.

SIGNATURE OF STUDY PARTICIPANT

Name of Participant

Signature of Participant

Date

SIGNATURE OF PERSON OBTAINING CONSENT

Name of Person Obtaining Consent

Contact Number

Signature of Person Obtaining Consent

Date

Protocol ID:IRB#13--001788 UCLA IRB Approved Approval Date: 1/8/2014 Through: 1/7/2017 Committee: North General IRB

ANEXO G

Roteiro para entrevista de participante
Experimento de percepção de vozes de falantes bilíngues
Ana Paula P. F. Engelbert

Participante # _____

Sexo: _____

Idade: _____

Local de nascimento: _____

Local onde cresceu (se diferente do de nascimento) _____

Idade que iniciou o aprendizado do inglês:

Tempo de estudo do inglês: _____

Experiência em países de língua inglesa:

() sim () não qual país? _____

Conhecimento em outra língua estrangeira:

() sim () não qual língua? _____

Autoavaliação de proficiência (de 0 a 100%): _____

Autoavaliação de grau de sotaque (de 0 a 100%): _____